Proceedings of the 2009 Winter Simulation Conference
M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin and R. G. Ingalls, eds.

OPTIMAL COMPUTING BUDGET ALLOCATION FOR CONSTRAINED OPTIMIZATION

Nugroho Artadi Pujowidianto
Loo Hay Lee

Dept. of Industrial and
Systems Engineering
National University of Singapore
10 Kent Ridge Crescent, 119260,
SINGAPORE

Chun-Hung Chen

Dept. of Systems Engineering
and Operations Research
George Mason University
4400 University Drive, MS 4A6
Fairfax, Virginia 22030, USA.

Chee Meng Yap

Dept. of Industrial and
Systems Engineering
National University of Singapore
10 Kent Ridge Crescent, 119260,
SINGAPORE

ABSTRACT

In this paper, we consider the problem of selecting the best design from a discrete number of alternatives in the presence of a stochastic constraint via simulation experiments. The best design is the design with smallest mean of main objective among the feasible designs. The feasible designs are the designs of which constraint measure is below the constraint limit. The Optimal Computing Budget Allocation (OCBA) framework is used to tackle the problem. In this framework, we aim at maximizing the probability of correct selection given a computing budget by controlling the number of simulation replications. An asymptotically optimal allocation rule is derived. A comparison with Equal Allocation (EA) in the numerical experiments shows that the proposed allocation rule gains higher probability of correct selection.

1 INTRODUCTION

In many applications, it is necessary to select the best design among competing designs. We consider the case where the performance measures have to be evaluated via simulation experiments. The advantage of simulation is its ability in capturing the dynamic relationships between parts of the object of the study and the uncertainty factors of which closed-form analytical solution may not be available. However, simulation is computationally intensive. One way to address this issue is by reducing the variance of the simulation outputs. Law (2007) described various types of variance reduction techniques and pointed out extensive references. The limitation of this technique is the needs of a detailed understanding of the model as the technique is very context-specific. This motivates the use of selection procedures called as Ranking and Selection (R&S).

R&S procedures are statistical methods for selecting the best design or the optimal subset from a discrete number of alternatives. Bechhofer et al. (1995), Swisher et al. (2003), and Kim and Nelson (2003) provided excellent review of the R&S works. One of the approaches in R&S is based on indifference zone (IZ). In this procedure, a certain level of probability of correct selection is guaranteed. The best design has to be better than other designs by a certain value so that the decision maker would not become indifferent. Rinott (1978) used two-stage-IZ procedure for allocating the computing budget in selecting the best design based on a single performance measure. The number of simulation replications in the second stage is allocated based on the variance of the results in the first stage. An example of fully-sequential-IZ procedures can be found in Kim and Nelson (2001).

Recently, R&S procedures with asymptotically optimal allocation rule have been shown to improve the probability of correct selection. Chen et al. (2000) and Chen and Yücesan (2005) aimed at maximizing the probability of correctly selecting the best design under a budget constraint and then derived the asymptotically optimal solution. The framework is called as Optimal Computing Budget Allocation (OCBA). It performs better than the procedure by Rinott (1978) because it uses the information of both relative means and variance instead of variance only. The application of OCBA could be found in the work by Chen et al. (2003).

The OCBA framework has been used in several ways. Fu et al. (2007) relaxed the independence assumption in OCBA by considering correlated sampling of the design performance. Assuming the performance measure across designs are independent, Chen et al. (2008) expanded the applicability of OCBA by developing the allocation rule in the context of finding an optimal subset instead of a single best design. Both Fu et al. (2007) and Chen et al. (2008) selected the best design based on a single performance measure. There may also be problems with multiple performance measures. Lee et al. (2004) provided the derivation of OCBA for multi-objective problems. Its application can be found in the work by Chew et al. (2009).

In the works mentioned previously, the performance measures are not restricted by certain limits. In reality, there are some problems with stochastic constraints. Hospital appointment scheduling problem is an example. As the resources are limited, the hospital administrators commonly attempt to minimize the resources idle time. At the same time, there is a waiting time limit set by the hospital or the government that restricts the selection of alternatives with high waiting time. Converting these two performance measures into a single weighted objective may face difficulties in specifying the weight. Considering them as separate objectives may waste computing budget as less computing budget should be allocated to the infeasible designs. This situation also applies to other problems where there are two kinds of performance measures. One measure would act as the main objective which needs to be optimized while the other measures in form of service criteria just need to be any values below the constraint limit.

Several researchers have responded to the needs of R&S procedures in the presence of stochastic constraints. Andradóttir et al. (2005) used two phases to select the best feasible design in the presence of one stochastic constraint. In the first phase, the feasible design will be selected. The concept of indifference-zone in form of pre-specified target level and tolerance level is applied in determining the feasibility. In the second phase, the best design among the feasible designs will be selected. The determination of feasible designs becomes more complex in the case of multiple constraints. Batur and Kim (2005) proposed a procedure to accelerate the computation for identifying the feasible designs. This is done by first eliminating unacceptable designs using a screening procedure based on aggregated observations. Similar to Andradóttir et al. (2005), Szechtman and Yücesan (2008) considered one stochastic constraint. They proposed a procedure for the first phase, namely identifying all feasible designs based on large deviations theory. The normality assumption is thus not needed and the result can be applicable for general distributions.

This paper provides an alternative approach based on OCBA framework to tackle the problem of selecting the best design in a constrained optimization problem. Unlike the previous works on R&S with constraints, the proposed approach does not need to first identify all feasible designs correctly. In this case, the computing budget for ensuring the correct decision in identifying all feasible designs could be saved. This is similar to the work by Morrice and Butler (2006) which use multi-attribute utility (MAU) theory. They extended the work of Butler et al. (2001) by specifying zero value in the utility function for infeasible designs. The limitation is the extra effort in eliciting the right utility functions and the relative importance across the performance measures.

In the work by Chen et al. (2000), the allocation is determined based on the variance and the distance between the mean of main objective of the non-best design and that of the best design. It does not use the information of the constraint measure variance and the distance between the mean of the constraint measure and the constraint limit. This paper attempts to improve the probability of correct selection in a constrained optimization problem by incorporating the information related to the constraint measure.

The organization of the paper is as follows. In the next section, the formulation of the computing budget allocation problem is provided together with the assumptions made. Section 3 proposes the rule for allocating the number of simulation replications for a constrained optimization problem, referred as Optimal Computing Budget Allocation for Constrained Optimization (OCBA-CO). The performance of the proposed allocation rule in the numerical experiments is shown in section 4. Section 5 concludes the paper and provides future research directions.

2 PROBLEM FORMULATION

This section first provides the formulation of a constrained optimization problem with stochastic performance measures. It is followed by the assumptions used. The computing budget allocation problem and the definition of probability of correct selection are then introduced. The notations are defined when they first appear.

We consider the problem of selecting the best design from a discrete number of alternatives in the presence of stochastic constraints. The performance measures considered have to be evaluated using simulation. Let Θ be the search space, an arbitrary, huge, structure less but finite set. There are k number of designs in the search space where θ_i is the system design parameter vector for design i, $i = 1,2,...,k$. J_0 indicates the mean of the main objective while J_h indicates the mean of the constraint measure h, $h = 1,2,..., A$ as there are A stochastic constraints. The mean of the main objective and each constraint measure are the expectation of L_h, the sample performance measure, a function of θ_i and ξ, the random vector representing the uncertain factors. σ_{0i} indicates the variance of the main objective value, ${\sigma_{0i}}^2 = Var(L_0(\theta_i, \xi))$, while $\sigma_{hi,h\neq 0}$ indicates the variance of the constraint measure h value, ${\sigma_{hi,h\neq 0}}^2 = Var(L_{h,h\neq 0}(\theta_i, \xi))$. A design is feasible if all constraint measures satisfy their constraint limit, c_h. The best design, θ_b is the design with smallest mean of main objective among the feasible designs. Therefore, the constrained optimization problem could be formulated as the following:

$$\min_{\theta_i \in \Theta} J_0(\theta_i) \equiv E[L_0(\theta_i, \xi)] \text{ subject to } J_{h,h\neq 0}(\theta_i) \equiv E[L_{h,h\neq 0}(\theta_i, \xi)] \leq c_h . \tag{1}$$

2.1 Assumptions

- $E[L_h(\theta_i, \xi)]$ can be estimated by the sample mean performance measure, namely $\bar{J}_{hi} = \frac{1}{N_i}\sum_{j=1}^{N_i} L_h(\theta_i, \xi_{ij})$. ξ_{ij} is the j-th simulation replication of the random vector that represents uncertain factors while N_i is the number of simulation replications for design i.
- The sample mean, $\bar{J}_{hi}$ follows normal distribution, $\bar{J}_{hi} \sim N\left(J_{hi}, \frac{\sigma_{hi}^2}{N_i}\right)$.
- As an initial attempt to develop the framework, it is assumed that $A = 1$ indicating that there is only one constraint, $c = c_1$.
- The simulation outputs from different replications are independent. In addition, the correlation of performance measures are not considered.

2.2 Computing Budget Allocation in a Constrained Optimization Problem

The main idea of OCBA is to control the number of simulation replications for each design, N_i so that the probability of correct selection, PCS is maximized given a total computing budget, T. Given that $N_i \geq 0$, the formulation of OCBA is

$$\max_{N_1,\ldots,N_k} PCS \text{ subject to: } N_1 + N_2 + \ldots + N_k = T . \tag{2}$$

Let $\theta_{bsample}$ is the best design based on sample mean performances while θ_b is the true best design. The probability of correct selection, PCS is the probability that the true best design is selected based on the sample mean performances, $PCS = P\{\theta_b = \theta_{bsample}\}$. For the true best design, θ_b to be selected in the simulation observation, the true best design must first remains feasible, namely the event of its constraint measure value satisfying the constraint limit, $(\bar{J}_{1b} \leq c)$. In addition, the true best design must be better than all other designs. The true best design is better than a non-best design in the observation if the non-best design is infeasible, $(\bar{J}_{1i} > c)$ or if the true best design has a smaller mean, $(\bar{J}_{0b} < \bar{J}_{0i})$ in the case where the non-best design is feasible, $(\bar{J}_{1i} \leq c)$. Therefore, for a constrained optimization problem, the PCS is defined as

$$PCS = P\left\{(\bar{J}_{1b} \leq c) \bigcap_{i=1,i\neq b}^{k} \left\{(\bar{J}_{1i} > c) \bigcup \left[(\bar{J}_{1i} \leq c) \bigcap (\bar{J}_{0b} < \bar{J}_{0i})\right]\right\}\right\}. \tag{3}$$

3 APPROXIMATE ASYMPTOTICALLY OPTIMAL SOLUTION

In our approach, the problem in (2) is approximated by

$$\max_{N_1,\ldots,N_k} APCS \text{ subject to: } N_1 + N_2 + \ldots + N_k = T . \tag{4}$$

In other word, the PCS is first replaced with an Approximate PCS $(APCS)$ so that the optimization problem of computing budget allocation can be solved analytically. The expression of $APCS$ and its derivation can be found in Lee et al. (2009).

Theorem 1 is the asymptotic solution which satisfies the Karush-Kuhn-Tucker (KKT) conditions. This result is referred as Optimal Computing Budget Allocation for Constrained Optimization (OCBA-CO). The detailed derivation and the sequential procedure for implementing the allocation rule can also be found in Lee et al. (2009).

Theorem 1 *As $T \to \infty$, the Approximate Probability of Correct Selection can be asymptotically maximized when the relationship between the number of simulation replications of two non-best designs is*

$$\frac{N_i}{N_j} = \left(\frac{\sigma_i / \delta_i}{\sigma_j / \delta_j} \right)^2 , \tag{5}$$

and the relationship between the number of simulation replications of the best design and non-best designs is

$$N_b = \sigma_{0b} \sqrt{\sum_{i \in \Theta_D} \left(\frac{N_i^2}{\sigma_{0i}^2} \right)} \quad if \quad \frac{\left(\sigma_{0b} \sqrt{\sum_{i \in \Theta_D} \left(\frac{N_i^2}{\sigma_{0i}^2} \right)} \right)}{(\sigma_b / \delta_b)^2} > \frac{N_i}{(\sigma_i / \delta_i)^2} , \tag{6}$$

$$\frac{N_b}{N_i} = \left(\frac{\sigma_b / \delta_b}{\sigma_i / \delta_i} \right)^2 \quad if \quad \frac{\left(\sigma_{0b} \sqrt{\sum_{i \in \Theta_D} \left(\frac{N_i^2}{\sigma_{0i}^2} \right)} \right)}{(\sigma_b / \delta_b)^2} < \frac{N_i}{(\sigma_i / \delta_i)^2} , \tag{7}$$

where $\Theta_D \equiv \{design \quad i \mid i \neq b, P(\bar{J}_{1i} \leq c) \geq P(\bar{J}_{0b} > \bar{J}_{0i})\}$ *and* $\Theta_F \equiv \{design \quad i \mid i \neq b, P(\bar{J}_{1i} \leq c) < P(\bar{J}_{0b} > \bar{J}_{0i})\}$. $\sigma_b / \delta_b = \sigma_{1b} / (J_{1b} - c)$. $\sigma_i / \delta_i = \sigma_{0i} / (J_{0i} - J_{0b})$ *if* $i \in \Theta_D$ *or* $\sigma_i / \delta_i = \sigma_{1i} / (J_{1i} - c)$ *if* $i \in \Theta_F$. *Similarly,* $\sigma_j / \delta_j = \sigma_{0j} / (J_{0j} - J_{0b})$ *if* $j \in \Theta_D$ *or* $\sigma_j / \delta_j = \sigma_{1j} / (J_{1j} - c)$ *if* $j \in \Theta_F$.

4 NUMERICAL EXPERIMENTS

In the numerical experiments, the performance of the OCBA rule for Constrained Optimization (OCBA-CO) is compared with the performance of Equal Allocation (EA). In EA, all designs are simulated equally. In this case, each alternative obtains T / k computing budget, allocated in one stage.

A simple minimization problem scenario is considered. A simple scenario is purposely used so that the true mean and variance are known. Therefore the true best design could be determined without simulation. This would enable the comparison of probability of correct selection of each rule by dividing the number of the true best design selected as the best design by the number of trials. 11 designs are considered. The mean of the main objective and the constraint measure for each design is a follows:

$$J_{0i} = i, i = 1,2,...,11, \tag{8}$$

$$J_{1i} = 12 - i, i = 1,2,...,11 . \tag{9}$$

One constraint limit value is used, $c = 5.5$, indicating that 40% of non-best designs are feasible. The true best design in this case is design 7. Equal variance is used with $\sigma_{0i} = \sigma_{1i} = 2$. The initial number of replications allocated is 10 for each alternative. The increment in each iteration is $k * 2$ which would be divided among the alternatives. 10,000 trials are conducted to compute the probability of correct selection, PCS , for each rule. For comparison purpose, four different values of PCS are used.

Table 1 shows the result of the numerical experiments. The computing budget required to reach each value of PCS using OCBA-CO and EA is recorded. The savings gained from using OCBA-CO over EA is represented by the speedup factor. The speedup factor is given by $T_{EA}/T_{OCBA-CO}$, the ratio between the total computing budget needed to reach the value of PCS using EA and that of OCBA-CO. For instance, EA requires 1,144 simulation replications in order to select the true best design correctly in 9,900 out of the 10,000 trials while OCBA-CO only needs 330 simulation replications. Therefore, OCBA-CO is 3.47 times faster than EA in reaching 99% PCS.

Table 1: The speedup factor gained from using OCBA-CO.

PCS	OCBA-CO	EA	Speedup Factor
90%	198	506	2.56
95%	220	682	3.10
97.5%	264	902	3.42
99%	330	1,144	3.47

It is shown that OCBA-CO performs better than EA. The reason is that EA does not allocate more simulation replications to the true best design and the non-best design with higher chance of being incorrectly selected as the best design based on the simulation output. In addition, EA does not consider the variance of the performance measures in allocating the computing budget. As the value of PCS increases, the speedup factor gained from using OCBA-CO instead of EA becomes larger. This indicates that performance of OCBA-CO is even more efficient when high PCS is required.

5 CONCLUSIONS

The problem of determining the number of simulation replications for each design in selecting the best design in the presence of one stochastic constraint is formulated as an optimization model. The objective is to maximize the probability of correct selection, PCS given a computing budget. The PCS is defined and an asymptotically optimal allocation rule which maximizes the approximate term of PCS is derived. The numerical results show that the Optimal Computing Budget Allocation procedure for Constrained Optimization (OCBA-CO) performs better than Equal Allocation (EA). Although the algorithm is based on asymptotic condition, it performs well with limited computing budget.

There are several future research directions. First, the correlation between the main objective and the constraint measure needs to be considered. In addition, the allocation rule needs to be extended to include multiple constraints. The third direction is to find an optimal subset instead of a single best design to provide the ability of screening. This could lead to the integration of OCBA-CO with a suitable search algorithm for a complete simulation optimization procedure.

REFERENCES

Andradóttir, S., D. Goldsman, and S.-H. Kim. 2005. Finding the best in the presence of a stochastic constraint. In *Proceedings of the 2005 Winter Simulation Conference*, ed. M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 732–738. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Batur, D. and S.-H. Kim. 2005. Procedures for feasibility detection in the presence of multiple constraints. In *Proceedings of the 2005 Winter Simulation Conference*, ed. M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 692–698. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Bechhofer, R. E., T. J. Santner, and D. M. Goldsman. 1995. *Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons*, New York: John Wiley & Sons.

Butler, J., D. J. Morrice, and P. W. Mullarkey. 2001. A multiple attribute utility theory approach to ranking and selection. *Management Science* 47:800-816.

Chen, C. H., K. Donohue, E. Yücesan, and J. Lin. 2003. Optimal computing budget allocation for Monte Carlo simulation with application to product design. *Simulation Modelling Practice and Theory* 11:57-74.

Chen, C. H., D. H. He, M. Fu, and L. H. Lee. 2008. Efficient simulation budget allocation for selecting an optimal subset. *INFORMS Journal on Computing* 20:579-595.

Chen, C. H., J. Lin, E. Yücesan, and S. E. Chick. 2000. Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems: Theory and Applications* 10:251-270.

Chen, C. H. and E. Yücesan. 2005. An alternative simulation budget allocation scheme for efficient simulation. *International Journal of Simulation and Process Modeling* 1:49-57.

Chew, E. P., L. H. Lee, S. Teng, and C. H. Koh. 2009. Differentiated service inventory optimization using nested partitions and MOCBA. *Computers and Operations Research* 36:1703-1710.

Fu, M. C., J. H. Hu, C. H. Chen, and X. Xiong. 2007. Simulation allocation for determining the best design in the presence of correlated sampling. *INFORMS Journal on Computing* 19:101-111.

Kim, S.-H. and B. L. Nelson. 2001. A fully sequential procedure for indifference-zone selection in simulation. *ACM Transactions on Modeling and Computer Simulation* 11:251-273.

Kim, S.-H. and B. L. Nelson. 2003. Selecting the best system: theory and methods. In *Proceedings of the 2003 Winter Simulation Conference*, ed. S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, 101-112. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Law, A.M. 2007. *Simulation modeling and analysis*. 4th ed. New York: McGraw-Hill, Inc.

Lee, L. H., E. P. Chew, S. Teng, and D. Goldsman. 2004. Optimal computing budget allocation for multi-objective simulation models. In *Proceedings of the 2004 Winter Simulation Conference,* ed. R. G. Ingalls, M. D. Rosetti, J. S. Smith, and B. A. Peters, 586-594. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Lee, L. H., N. A. Pujowidianto, L.-W. Li, C. H. Chen, and C. M. Yap. 2009. Asymptotic simulation budget allocation for selecting the best design in the presence of stochastic constraints. Working Paper, Department of Industrial and Systems Engineering, National University of Singapore.

Morrice, D. J. and J. C. Butler. 2006. Ranking and selection with multiple "targets". In *Proceedings of the 2006 Winter Simulation Conference*, ed. L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 222–230. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Rinott, Y. 1978. On two-stage selection procedures and related probability-inequalities. *Communications in Statistics–Theory and Methods* A7:799-811.

Swisher, J. R., S. H. Jacobson, and E. Yücesan. 2003. Discrete-event simulation optimization using ranking, selection, and multiple comparison procedures: a survey. *ACM Transactions on Modeling and Computer Simulation* 13:134-154.

Szechtman, R. and E. Yücesan. 2008. A new perspective on feasibility determination. In *Proceedings of the 2008 Winter Simulation Conference*, ed. S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler, 273–280. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

NUGROHO ARTADI PUJOWIDIANTO is a Ph.D. candidate in the Department of Industrial and Systems Engineering, National University of Singapore. He received his B.Eng. (Mechanical Engineering) degree with a specialization in Manufacturing and Systems Engineering from Nanyang Technological University, Singapore in 2006. His research interest is in simulation optimization and its application in health care. His current work is on the issue of computing budget allocation for a constrained optimization problem with uncertainty. His email address is <nugroho@nus.edu.sg>.

LOO HAY LEE is an Associate Professor in the Department of Industrial and Systems Engineering, National University of Singapore. He received his B.S. (Electrical Engineering) degree from the National Taiwan University in 1992 and his S. M. and Ph.D. degrees in 1994 and 1997 from Harvard University. He is currently a senior member of IEEE, a committee member of ORSS, and a member of INFORMS. His research interests include production planning and control, logistics and vehicle routing, supply chain modeling, simulation-based optimization, and evolutionary computation. His email address is <iseleelh@nus.edu.sg>.

CHUN-HUNG CHEN is a Professor of Systems Engineering and Operations Research at George Mason University. He received his Ph.D. from Harvard University in 1994. His research interests are mainly in development of very efficient methodology for simulation and optimization and its applications. Dr. Chen has served as Co-Editor of the Proceedings of the 2002 Winter Simulation Conference and Program Co-Chair for 2007 Informs Simulation Society Workshop. He is currently an associate editor of IEEE Transactions on Automatic Control, area editor of Journal of Simulation Modeling Practice and Theory, associate editor of International Journal of Simulation and Process Modeling, and simulation department editor for IIE Transactions. His email address is <cchen9@gmu.edu>.

CHEE MENG YAP is a Senior Lecturer in the Department of Industrial and Systems Engineering, National University of Singapore. He received his B. Eng. (First Class Honours) in Civil Engineering from National University of Singapore in 1989 and his M.S. in Industrial Engineering and Ph.D. in Engineering Management in 1989 and 1992 from University of Pittsburgh. In addition, he holds the Chartered Financial Analyst designation. His research interests include R&D investment and firm value, and operations management in the healthcare industry. His email address is <iseyapcm@nus.edu.sg>.