BETTER-THAN-OPTIMAL SIMULATION RUN ALLOCATION?

Chun-Hung Chen Donghai He

Dept. of Systems Engineering & Operations Research George Mason University 4400 University Drive, MS 4A6 Fairfax, VA 22030, U.S.A.

ABSTRACT

Simulation is a popular tool for decision making. However, simulation efficiency is still a big concern particularly when multiple system designs must be simulated in order to find a best design. Simulation run allocation has emerged as an important research topic for simulation efficiency improvement. By allocating simulation runs in a more intelligent way, the total simulation time can be dramatically reduced. In this paper we develop a new simulation run allocation scheme. We compare the new approach with several different approaches. One benchmark approach assumes that the means and variances for all designs are known so that the theoretically optimal allocation can be found. It is interesting to observe that an approximation approach called OCBA does better than this theoretically optimal allocation. Moreover, a randomized version of OCBA may outperform OCBA in some cases.

1 INTRODUCTION

Simulation is a popular tool for analyzing systems and evaluating decision problems since real situations rarely satisfy the assumptions of analytical models. While simulation has many advantages for modeling complex systems, efficiency is still a significant concern when conducting simulation experiments (Law and Kelton 2000). To obtain a good statistical estimate for a design decision, a large number of simulation samples or replications is usually required for each design alternative. If the accuracy requirement is high, and if the total number of designs in a decision problem is large, then the total simulation cost can easily become prohibitively high.

The problem we consider is that of selecting the best design among a finite number of choices, where the performance of each design must be estimated with some uncertainty, specifically through stochastic sampling. The primary context is that of simulation, where the goal would be to determine the best allocation of simulation replicaEnver Yücesan

Technology Management Area INSEAD Boulevard de Constance 77305 Fontainebleau Cedex, FRANCE

tions among competing designs. This problem setting falls under the well-established branch of statistics known as ranking and selection and/or multiple comparison procedures. In the context of simulation, Goldsman and Nelson (1998) provide a nice overview of this field. The ranking and selection algorithms determine the number of simulation replications required for each design in order to guarantee a pre-specified level of correct selection, whereas multiple comparison procedures provide confidence intervals on estimated performance differences between systems. More recently, Chen et al. (1997, 2000) and Chick and Inoue (2001ab) have approached the problem from the perspective of allocating a fixed number of simulation replications in order to maximize the probability of correct selection. Chen has called this optimal computing budget allocation (OCBA).

Intuitively, to ensure a high probability of correctly selecting a good design (the so-called probability of correct selection, $P\{CS\}$), a larger portion of the computing budget should be allocated to those designs that are critical in the process of identifying good designs. In other words, a larger number of simulations must be conducted with those critical designs in order to reduce estimator variance. On the other hand, limited computational effort should be expended on non-critical designs that have little effect on identifying the good designs even if they have large variances. Overall simulation efficiency is improved as less computational effort is spent on simulating non-critical designs and more is spent on critical designs. Ideally, one would like to allocate simulation runs to designs in a way that maximizes $P\{CS\}$ within a given computing budget. However, solving such an optimal run allocation problem is a big challenge because i) there is no closed-form expression for $P{CS}$; ii) $P{CS}$ is a function of the means and variances of all designs which are unknown; and iii) a solution should be found efficiently. Otherwise the benefit of efficient run allocation will be lost.

In this paper, we develop a new approach based on OCBA. We also compare this new approach with a few

different approaches including the theoretically optimal allocation, equal allocation, a Rinott-type approach (1978), and OCBA given by Chen et al. (2000). The theoretically optimal allocation, which serves as a benchmark for comparison, assumes that the means and variances for all designs are known so that the optimal allocation for maximizing $P\{CS\}$ can be found. It is interesting to observe that OCBA and the newly developed approach perform better than the theoretically optimal allocation.

The paper is organized as follows: In the next section, we define the notation and the simulation run allocation problem. Section 3 presents our new simulation run allocation approach. Section 4 gives a brief description for all other approaches we are comparing in the paper. Numerical experiments are given in Section 5. Section 6 concludes the paper.

2 PROBLEM STATEMENT

Our goal is to select a design associated with the smallest mean performance measure among k alternative designs with unequal and possibly unknown variances. We consider terminating (finite-horizon) simulations in this paper. (Our approach is equally applicable to steady-state simulations where we need N approximately independent samples rather than N independent simulation replications. In that case, the batch means method can be applied to approximate the independence of the samples.) We further assume that the simulation output is independent from replication to replication. The sampling across designs is also independent. Suppose that the computing budget is limited. Denote by

- X_{ij} : the *j*-th independent and identically distributed (*i.i.d.*) sample of the performance measure from design *i*. To simplify the illustration, X_{ij} is assumed to be normally distributed in this paper.
- N_i : the number of simulation runs for design *i*,
- \overline{X}_i : the sample average of the simulation output for

design *i*;
$$\overline{X}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} X_{ij}$$

- S_i^2 : the sample variance of the simulation output for design *i*,
- μ_i : the unknown mean performance measure; $\mu_i = E[X_{ij}]$,
- σ_i^2 : the variance for design *i*, i.e., $\sigma_i^2 = \text{Var}(X_{ij})$. In practice, σ_i^2 is unknown beforehand and so is approximated by sample variance.
- b: the design with the smallest sample mean performance; $b = \arg \min_{i} \{\overline{X}_i\}.$

$$\delta_{b,i} \equiv \overline{X}_b - \overline{X}_i.$$

While the design with the smallest sample mean (design b) is usually selected, design b is not necessarily the one with the smallest unknown mean performance. Correct selection (CS) is therefore defined as the event that design b is actually the best design (i.e., with the smallest population mean, hence the true best design). In this paper, we wish to maximize the probability of correctly identifying the true best design, $P{CS}$, with a given computing budget T. Assume that the computation cost for each run is roughly the same across different designs. The computation cost can then be approximated by $N_1 + N_2 + \dots + N_k$, the total number of runs. In the numerical experiments given in Section 4, we will compare the probability of correct selection $P\{CS\}$ that each procedure can obtain with a given computing budget. The achieved $P{CS}$ serves as a measure of the effectiveness for each approach.

3 A NEW COMPUTING BUDGET ALLOCATION APPROACH

In this section, we present a new computing budget allocation scheme called Randomized OCBA (ROCBA). This is developed based on the OCBA algorithm due to Chen et al. (2000). OCBA allocates simulation runs by considering the following optimization problem:

$$\max_{N_1, \dots, N_k} P\{\text{CS}\}$$

s.t. $N_1 + N_2 + \dots + N_k = T.$ (1)

Under a Bayesian model, OCBA approximates $P\{CS\}$ using the Bonferroni inequality and offers an asymptotic solution to this approximation. While the run allocation given by OCBA is not an optimal allocation when the simulation budget is finite, the numerical testing demonstrates that OCBA is a very efficient approach and can dramatically reduce simulation time. In particular, OCBA allocates simulation runs according to:

$$\frac{N_i}{N_j} = \left(\frac{\sigma_i/\delta_{b,i}}{\sigma_j/\delta_{b,j}}\right)^2,$$

i, j $\in \{1, 2, ..., k\}$, and $i \neq j \neq b$, (2)

$$N_b = \sigma_b \sqrt{\sum_{i=1, i \neq b}^k \frac{N_i^2}{\sigma_i^2}}$$
(3)

Based on OCBA, we develop a new allocation scheme ROCBA. While OCBA is simply an asymptotic solution to an approximation, the solution can be calculated easily and is proven to be effective. On the other hand, we know there is a potential to further improve the performance of OCBA. But the original budget allocation problem (1) is extremely difficult to solve. It may not be worthy to insist on spending much time on finding an optimal solution. Instead of trying to solve such a hard problem, we consider a quick alternative.

Our idea is to have the asymptotic solution from OCBA to form the basis for simulation budget allocation, but with some minor random perturbation. We first apply OCBA to find the asymptotic solution of (N_1, N_2, \dots, N_k) using (2) and (3). Thus, the proportion of the budget allocated to design *i*, $\alpha_i = N_i / (N_1 + N_2 + \dots + N_k)$, can also be calculated. We know that α_i obtained this way is not a best solution; but we believe it is close to the best solution. Our approach is to make some perturbation on the budget allocation proportion α_i by multiplying it by a random number, $(1 + r_i)$, where r_i is uniformly distributed between $-\delta$ and δ . Then we use the ratio of $(1 + r_1)^* \alpha_1 : (1 + r_2)^* \alpha_2 : \dots : (1 + r_k)^* \alpha_k$ to determine the budget allocation.

Specifically, the original ratio for budget allocation obtained from OCBA is $\alpha_1 : \alpha_2 : \dots : \alpha_k$; the one for our new approach ROCBA is $(1 + r_1)^* \alpha_1 : (1 + r_2)^* \alpha_2 : \dots : (1 + r_k)^* \alpha_k$. With a given computing budget, this ratio is used to determine the allocation of simulation runs. In this setting, r_i can be viewed as the random proportion with which we perturb α_i from the value calculated using OCBA.

It is certain that the value of δ will have significant impact on the performance of the new algorithm ROCBA. We don't want the perturbed allocation to be too much different from the allocation given by OCBA. As OCBA can provide a very good allocation, δ must be a very small positive number. In this paper, we set $\delta = 0.035$. A good selection for the value of δ warrants more research.

4 DIFFERENT ALLOCATION PROCEDURES

In this section, we test our new allocation approach and compare it with several different allocation procedures through a series of numerical experiments. Among them, equal allocation represents the most straightforward way of conducting simulation experiments. OCBA is developed based on a Bayesian framework and intends to optimize an approximation of $P\{CS\}$. Finally, Rinott is highly popular in simulation literature.

To have a better comparison, we put all tested procedures in the same setting of sequential sampling. Initially, n_0 simulation runs for each of k designs are performed to get some information about the performance of each design during the first stage. As simulation proceeds, the sample means and sample variances of all designs are computed from the data already collected up to that stage. According to this collected simulation output, an incremental computing budget, Δ , is allocated based on different allocation approaches at each stage. The procedure is continued until the total budget *T* is exhausted. Then the whole procedure is repeated for 1,000,000 times. We estimate $P\{CS\}$ by counting the number of times we successfully find the true best design out of 1,000,000 independent applications of each selection procedure. $P\{CS\}$ is then obtained by dividing this number by 1,000,000, representing the correct selection frequency. A million macro replications guarantee a standard error of the $P\{CS\}$ estimate under 0.001 or 0.1%. The $P\{CS\}$ obtained from each different procedure will serve as a measurement of its effectiveness.

In addition to different allocation approaches, we include a theoretic optimal allocation as a benchmark in our comparison. We briefly summarize the compared allocation procedures as follows.

4.1 Theoretically Optimal Allocation (TOA)

To compute the highest achievable $P\{CS\}$ under perfect information, we assume that the means and variances for all designs are known in this allocation. Thus $P\{CS\}$ can be calculated (or estimated through Monte Carlo simulation) if the value of (N_1, N_2, \dots, N_k) is given.

$$P\{CS\} = P\{\overline{X}_{best design} < \overline{X}_i, \text{ for all } i \text{ where design } i \text{ not the best } \}.$$

Since the total computing budget, *T*, considered in this paper is not big, we can evaluate $P\{CS\}$ for all possible combinations of (N_1, N_2, \dots, N_k) with a constraint that $N_1 + N_2 + \dots + N_k = T$. Then the maximum $P\{CS\}$ and the corresponding (N_1, N_2, \dots, N_k) can be determined. Such a maximum $P\{CS\}$ will serve as a benchmark for comparison.

It is unrealistic to assume that the means and variances of all designs are known prior to performing simulations. For all other compared approaches, we do not make such an assumption. The sample means and sample variances of all designs are computed from the collected data and then used to determine the run allocations.

4.2 Equal Allocation (Equal)

This is the simplest way to conduct simulation experiments and has been widely applied. The simulation budget is equally allocated to all designs. Namely, all designs are equally simulated with $N_i = T/k$ for each *i*.

4.3 Proportional To Variance (PTV)

The two-stage procedure of Rinott (1978) has been widely applied in the simulation literature (Law and Kelton 2000). See Bechhofer et al. (1995) for a systematic discussion of two-stage procedures. In the first stage, all designs are simulated for n_0 samples. Based on the sample variance estimate (S_i^2) obtained from the first stage, the number of

additional simulation samples for each design in the second stage is determined by:

$$N_i = \max(0, \lceil (S_i^2 h^2 / d^2 \rceil - n_0), \text{ for } i = 1, 2, ..., k, \quad (4)$$

where $\lceil \bullet \rceil$ is the integer "round-up" function, *d* is the indifference zone, h is a constant which solves Rinott's integral (h can also be found from the tables in Wilcox 1984). The major drawback is that only the information on variances is used when determining the simulation allocation, while the OCBA algorithm utilizes the information on both means and variances. As a result, the performance of Rinott's procedure is not as good as OCBA. We do, however, include it in our testing due to its popularity in the simulation literature.

To put this into the setting of our sequential sampling framework, we make some modifications to the standard two-stage Rinott procedure. The runs allocation in (4) implies that N_i is proportional to the estimated sample variances. In this procedure, the available computing budget is allocated in a way that N_i is proportional to the estimated sample variances.

4.4 Optimal Computing **Budget Allocation (OCBA)**

Under a Bayesian model, OCBA approximates $P\{CS\}$ using the Bonferroni inequality and offers an asymptotic solution to this approximation. The asymptotic solution is given in (2) and (3). In this approach, simulation runs are allocated based on (2) and (3).

NUMERICAL TESTING 5

The numerical experiments include a series of generic tests. In all the numerical illustrations, we estimate $P{CS}$ by measuring the relative frequency of the event that we successfully find the true best design (design 1 in this example) out of 1,000,000 independent applications of each selection procedure. The P{CS} obtained from each different procedure will serve as a measure of its effectiveness. We have set $n_0 = 10$ and $\Delta = 20$.

5.1 Experiment 1: Three Designs

This is a special case where the best design has zero variance. There are three design alternatives:

- $X_{1j} \sim N(0, 0^2)$, $X_{2j} \sim N(0.4, 1.5^2)$, and $X_{3j} \sim N(0.4, 3^2)$.

Suppose that the total computing budget $T = N_1 + N_2 + N_3$ = 100. In this case.

$$P\{CS\} = Pr\{X_{1}(N_{1}) < X_{2}(N_{2}) \text{ and } X_{1}(N_{1}) < X_{3}(N_{3})\}$$

= Pr{ 0 < $\overline{X}_{2}(N_{2})$ and 0 < $\overline{X}_{3}(N_{3})$ }
= Pr{ $\overline{X}_{2}(N_{2}) > 0$ } Pr{ $\overline{X}_{3}(N_{3}) > 0$ }
= $\Phi(\frac{0.4}{1.5/\sqrt{N_{2}}}) \Phi(\frac{0.4}{3.0/\sqrt{N_{3}}})$

Since the variance of design 1 is zero, we know that in the theoretically optimal allocation, we should not allocate any runs to design 1 as it will not further reduce its estimation variance. We should allocate the limited computing budget to Designs 2 and 3 only. Thus we can easily evaluate $P\{CS\}$ for all 101 combinations with the constraint N_2 $+ N_3 = 100$. The best of these 101 different allocations is when $N_2 = 38$ and $N_3 = 62$, with $P\{CS\} = 81.1\%$.

After determining that the theoretically optimal allocation can achieve $P{CS} = 81.1\%$, we use it as a benchmark to compare the performance of all other approaches. For all other approaches, we assume that the means and variances for all designs are unknown. The sample means and sample variances of all designs are computed from the collected data and then used to determine the run allocations. Table 1 shows the test results using different allocation procedures.

Table 1: Performance Comparison of Different Simulation Run Allocation Procedures in Experiment 1.

	TOA	Equal	PTV	OCBA	ROCBA
			(Rinott)		
<i>P</i> {CS}	81.0%	74.0%	76.1%	88.6%	88.5%

From Table 1, we see that TOA performs better than Equal and PTV. However, it is interesting to observe that both OCBA and ROCBA perform better than TOA. This seems counter intuitive. We conduct further testing and discuss them in the next section.

5.2 Experiment 2: Five Designs

We expand the special case considered in Experiment 1 to include 5 design alternatives:

- $X_{1i} \sim N(0, 0^2),$

- $X_{1j} \sim N(0, 0^{\circ}),$ $X_{2j} \sim N(1, 9^2),$ $X_{3j} \sim N(2, 9^2),$ $X_{4j} \sim N(3, 9^2),$ and $X_{5j} \sim N(4, 9^2).$

We set the total computing budget $T = N_1 + N_2 + N_3 + N_4 + N_4$ $N_5 = 600$. Because of zero variance for Design 1, we can again simplify the computation of $P\{CS\}$ for TOA.

$$P\{CS\} = \Pr\{\overline{X}_{1}(N_{1}) < \overline{X}_{i}(N_{i}), \text{ for } i = 2, 3, 4, \text{ and } 5\}$$

= $\Pr\{\overline{X}_{2}(N_{2}) > 0\} \Pr\{\overline{X}_{3}(N_{3}) > 0\}$
 $\Pr\{\overline{X}_{4}(N_{4}) > 0\} \Pr\{\overline{X}_{5}(N_{5}) > 0\}$
= $\Phi(\frac{1}{\frac{9}{\sqrt{N_{2}}}}) \Phi(\frac{2}{\frac{9}{\sqrt{N_{3}}}}) \Phi(\frac{3}{\frac{9}{\sqrt{N_{4}}}}) \Phi(\frac{4}{\frac{9}{\sqrt{N_{5}}}})$

While the number of combinations that $N_2 + N_3 + N_4 + N_5$ = 600 is large, it is still feasible to evaluate P{CS} for all combinations. The best allocation is when $N_2 = 348$, $N_3 =$ 134, $N_4 = 72$, and $N_5 = 46$, $P\{CS\} = 97.2\%$.

As in Experiment 1, we assume the means and variances for all designs are unknown for all other compared approaches. Table 2 shows the test results using different allocation procedures.

Table 2: Performance Comparison of Different Simulation Run Allocation Procedures in Experiment 2.

	TOA	Equal	PTV	OCBA	ROCBA
			(Rinott)		
<i>P</i> {CS}	97.2%	88.2%	91.2%	99.2%	99.3%

From Table 2, we see that TOA performs better than Equal and PTV. Again we observe that both OCBA and ROCBA outperform TOA. As in Experiment 1, the performances of OCBA and ROCBA are quite close.

5.3 Experiment 3: Five Designs with Non-Zero Variance

This example is an extension of Experiment 2. The variance of Design 1 is set to be the same as that of other designs, which is non-zero. Thus the 5 design alternatives are:

- $X_{1j} \sim N(0, 9^2), X_{2j} \sim N(1, 9^2), X_{3j} \sim N(1, 9^2), X_{3j} \sim N(2, 9^2), X_{3$
- $X_{4j} \sim N(3, 9^2)$, and
- $X_{5i} \sim N(4, 9^2).$

The total computing budget $T = N_1 + N_2 + N_3 + N_4 + N_4$ N_5 is also 600. In this case, we evaluate $P\{CS\}$ using Monte Carlo simulation for all combinations. The best allocation is when $N_1 = 216$, $N_2 = 174$, $N_3 = 113$, $N_4 = 59$, and $N_5 = 38$, $P\{CS\} = 83.6\%$. Table 3 shows the test results using different allocation procedures.

Once again, we see that TOA performs better than Equal and PTV and OCBA beats TOA. In this case ROCBA's performance is even better than OCBA.

Table 3: Performance Comparison of Different Simulation Run Allocation Procedures in Experiment 3.

	TOA	Equal	PTV	OCBA	ROCBA
			(Rinott)		
<i>P</i> {CS}	83.6%	78.6%	78.7%	84.8%	86.1%

6 **CONCLUDING REMARKS**

In this paper, we present a new simulation run allocation approach called ROCBA. ROCBA determines the run allocation by randomly perturbing the asymptotic solution obtained from OCBA. We test ROCBA and compare it with several different allocation procedures through a series of numerical experiments. One of the compared approaches assumes that the means and variances for all designs are known so that the theoretically optimal allocation can be found. We also compare it with equal allocation and a revised Rinott procedure. We had some interesting observations:

- OCBA and ROCBA perform better than Theoretically Optimal Allocation. This seems counter intuitive. Note that means and variances are assumed known in TOA, while it is not for OCBA/ROCBA. OCBA/ROCBA use only information of sample means and sample variances. Our conjecture is that OCBA/ROCBA is a dynamic allocation scheme, whereas TOA is static allocation. Using information of sample means and sample variances can become advantageous in some cases, because it contains some valuable information about what has been sampled. Thus a good simulation approach can utilize this additional information to do a better simulation run allocation, and beats the performance of the TOA.
- $P\{CS\}$ is not sensitive to the run allocations deter-. mined by OCBA. ROCBA is actually a deviation of OCBA by randomly perturbing its allocation. In the numerical testing, we found that the performances of OCBA and ROCBA are very close. The allocation determined by OCBA seems quite robust with some minor level of perturbation.
- ROCBA may beat OCBA in some cases. It is interesting to see that ROCBA performs slightly better than OCBA in Experiment 3. It is certain that the performance of ROCBA critically depends on the selection of δ . More studies are needed to give further implications.

REFERENCES

- Bechhofer R. E., T. J. Santner, and D. M. Goldsman, Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons, John Wiley & Sons, Inc., 1995.
- Chen, H. C., C. H. Chen, L. Dai, E. Yücesan, "New Development of Optimal Computing Budget Allocation for

Discrete Event Simulation," Proceedings of the 1997 Winter Simulation Conference, ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson, 334-341, 1997.

- Chen, C. H., J. Lin, E. Yücesan, and S. E. Chick, "Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization," Journal of Discrete Event Dynamic Systems: Theory and Applications, Vol. 10, pp. 251-270, 2000.
- Chick, S.E. and K. Inoue, "New Two-Stage and Sequential Procedures for Selecting the Best Simulated System," Operations Research, Vol. 49, pp. 732-743, 2001a.
- Chick, S.E. and K. Inoue, "New Procedures to Select the Best Simulated System Using Common Random Numbers," Management Science, Vol. 47, No.8, pp. 1133-1149, August 2001b.
- Goldsman, D. and B. L. Nelson, "Comparing Systems via Simulation," Chapter 8 in J. Banks, ed. Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice, John Wiley & Sons, New York. pp. 273-306, 1998.
- Law, A. M. and W. D. Kelton, Simulation Modeling & Analysis. McGraw-Hill, Inc, 2000.
- Rinott, Y., "On Two-stage Selection Procedures and Related Probability Inequalities," Communications in Statistics A7, 799-811, 1978.
- Wilcox, R. R. 1984. A Table For Rinott's Selection Procedure. Journal of Quality Technology, 16. 2, 97-100.

ACKNOWLEDGMENTS

This work has been supported in part by NSF under Grants DMI-0002900 and DMI-0049062, by NASA Ames Research Center under Grant NAG-2-1565, by FAA under Grant 00-G-016, by Sandia National Laboratories under Contract BD-0618, and by George Mason University Research Foundation.

AUTHOR BIOGRAPHIES

CHUN-HUNG CHEN is an Associate Professor of Systems Engineering at George Mason University. He served as the Co-Editor of the *Proceedings of the 2002 Winter Simulation Conference*. 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 application to engineering design and air traffic management. He is a member of INFORMS and a senior member of IEEE. His email address is <cchen9@gmu.edu>.

DONGHAI HE is in the Department of Systems Engineering & Operations Research at George Mason University. He received his BS degree in Control Engineering at Dalian University of Technology in China. His research interests include optimization, object-oriented modeling and air traffic management. His email address is <<u>dhe@gmu.edu</u>>.

ENVER YÜCESAN is a Professor of Operations Management at INSEAD in France. He served as the Co-Editor of the Proceedings of the 2002 Winter Simulation Conference. He received his MS and PhD degrees in Operations Research at Cornell University, and his BS degree in Industrial Engineering at Purdue University. His research interests are in modeling methodology, experimental design, and webbased applications. He is a member of INFORMS and IIE. His email address is <enver.yucesan@insead.edu>.