

Received April 28, 2020, accepted May 13, 2020, date of publication May 18, 2020, date of current version June 1, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2995196

A Simulation Budget Allocation Procedure for Finding Both Extreme Designs Simultaneously in Discrete-Event Simulation

SEON HAN CHOI¹, (Member, IEEE), AND CHANGBEOM CHOI², (Member, IEEE)

¹Department of IT Convergence and Application Engineering, Pukyong National University, Busan 48513, South Korea

²School of Global Entrepreneurship and ICT, Handong Global University, Pohang 37554, South Korea

Corresponding author: Changbeom Choi (cbchoi@handong.edu)

This work was supported by the National Research Foundation of Korea (NRF) funded by the Korea Government (Ministry of Science and ICT) under Grant 2019R1G1A1098951.

ABSTRACT Although discrete-event simulation has been widely used in various engineering fields, its efficiency remains an issue. Ranking and selection (R&S) procedures can solve this efficiency problem by allocating a limited simulation budget intelligently. While the existing R&S procedures mostly aim to find only the best simulation input design, practitioners sometimes require the worst design as well to analyze systems requiring high reliability, such as military systems, municipal waste management, etc. Motivated by these practical needs, we propose a simulation budget allocation procedure for selecting both extreme designs simultaneously in the presence of large stochastic noise. To maximize the accuracy of the selections under a limited budget, the proposed procedure sequentially allocates a small budget and updates the simulation results such that they can be used as significant evidence for the correct selections. Our experimental results on benchmark and practical problems demonstrate improved efficiency compared to previous works. It is expected that the proposed procedure will be effectively utilized in the fields of the fourth industrial revolution, such as digital twins that demand quickly finding both extreme designs to maintain synchronization with the corresponding real systems.

INDEX TERMS Discrete-event simulation, stochastic system, ranking and selection, extreme designs selection, simulation-based optimization.

I. INTRODUCTION

Discrete-event simulation is widely used to analyze modern industrial complex systems, such as manufacturing [1], military [2], smart grid [3], telecommunications [4], and transportation [5]. The most significant advantage of the simulation is that, with just a few assumptions, it can accurately analyze complex systems that cannot be described using closed-form analytic models [6]. However, the efficiency of discrete-event simulation is still a concern. Typical simulation models have random variables or processes to represent the uncertainty of the real world, resulting in stochastic noise in the simulation output. Thus, the expected value of the output for a simulation input design (e.g., a particular configuration of decision variables) can only be estimated as the average of output samples obtained from independently

repeated simulations (i.e., simulation replications). As the noise increases, a greater number of replications are required per design to obtain an accurate estimate, and the total computational costs may be prohibitively expensive as the number of inputs increases. Moreover, with the higher complexity of recent industrial systems, the increasing cost per simulation run intensifies the efficiency problem. Consequently, how to allocate a simulation budget (i.e., a limited number of simulation replications) efficiently to find the desired input design has drawn considerable attention in the field of simulation [7].

If the objective is to select the best simulation input designs from a finite set of alternatives, ranking and selection (R&S), a well-established branch of statistics, can be an efficient solution [8]–[10]. The best design here is to minimize (or maximize in some cases) the expected value of the simulation output. Based on the ordinal optimization, R&S intelligently allocates a limited number of simulation replications to correctly find the best design depending on

The associate editor coordinating the review of this manuscript and approving it for publication was Donatella Darsena¹.

the statistical comparisons and inferences for the designs' simulation results. Thus far, many efficient R&S procedures have been developed. Based on the sequential allocation procedure, these procedures iterate the following two steps to find the best design correctly: 1) allocate a few replications depending on the statistical inference for the designs' simulation results, then 2) update the results via additional simulations. R&S procedures' distinctive feature lies in the allocation step; that is, how to allocate those few further replications to maximize the statistical evidence for the correct selection of the best design [11].

The two-stage indifference-zone (IZ) procedure [12] focuses on frequentist evidence to achieve a predefined lower bound guarantee of the probability of correct selection $P\{CS\}$. It allocates further replications to the IZ designs that have simulation results similar to the current best design. Unlike the IZ procedure, the optimal simulation budget allocation (OCBA) procedure [13] applies Bayesian posterior distribution to describe the evidence of the correct selection. The OCBA procedure allocates further replications using the derived optimal allocation rules to asymptotically maximize an approximated lower bound of $P\{CS\}$. Although it is typically more efficient than the IZ procedure [11], due to the nature of the allocation rules that use the sample mean and variance values, the OCBA procedure may be inefficient in the presence of large stochastic noise, where the sample mean has an inaccurate value [14]. Recently, the uncertainty evaluation (UE) procedure has been proposed to efficiently solve practical problems that have large noise [15]. To maximize $P\{CS\}$, the UE procedure allocates further replications heuristically depending on the uncertainty defined with p -values in the statistical hypothesis test. Compared to the OCBA procedure, it has high robustness against noise by considering the precision of the sample mean additionally.

Most previous R&S procedures aim to find only the best input design or subset. However, practitioners in engineering sometimes require the worst design as well as the best design. In other words, both extreme designs are desired. For example, when analyzing a battleship's decoy system for torpedo avoidance, various scenarios that the battleship may encounter should be considered. The survival rate of the battleship for each scenario design can be estimated over many simulation replications of the anti-torpedo combat simulator [16]. Given k scenario designs, practitioners are interested in not only the best design, but also the worst design that minimizes the survival rate: even if the survival rate for the best scenario is very high, the decoy system should be reconsidered if its worst-case scenario occurs frequently in practice. In another example, when investigating the layout of a municipal sewage treatment plant in a city [17], practitioners demanded to find the best and worst layout designs simultaneously to make reliable and suitable layout plans.

In addition to these examples, selecting both extreme designs simultaneously is useful for multiple-criteria decision-making (MCDM) problems, which are commonly encountered in practice. This is because the selected extreme

designs for each criterion can be used to construct the ideal and nadir points, providing insight to decision-makers. These designs are also of worth in best-worst scaling, a famous discrete-choice model that requires picking the highest- and lowest-utility items [18]. Several population-based optimization algorithms, such as the ant colony algorithm, genetic algorithm, and particle swarm optimization, require finding the best and worst designs among the population to increase its quality [19], [20]. For example, one of the selection schemes in a genetic algorithm is to push the elite design directly into the next population and eliminate the worst design. In summary, the philosophy of finding both extreme designs has tremendous potential.

However, little research exists on selecting the best and worst designs. Zhang *et al.* [21] proposed the OCBA_{bw} procedure based on the OCBA procedure, and Xiao *et al.* [22] developed the OCBA_{mn} procedure for selecting the best and worst subsets by extending the OCBA_{bw} procedure. In addition, Zhang *et al.* [23] generalized the problem of OCBA_{bw} and suggested the OCBA_{sr} procedure to select n ranked subsets. As mentioned previously, these OCBA-based procedures may be inefficient when large stochastic noise exists, such as simulation models for modern industrial systems. Meanwhile, without loss of generality, finding the worst design is mathematically the same as finding the best one; thus, we can select both extreme designs at the same time by using the existing R&S procedures (such as the UE procedure) twice. For example, dividing a given simulation budget in half, half can be used to find the best design and the other half to find the worst design. This method can be effective in specific cases where the budgetary need to find both designs is similar; otherwise, it is inefficient (see the experimental results of the hOCBA and hUE procedures in Table 2). In addition, because prior information for the simulation output of given designs is typically unknown before conducting the simulation, it can be difficult to decide how best to divide a budget.

In this paper, we propose an R&S procedure for selecting both extreme designs simultaneously. The objective of the proposed procedure is to maximize $P\{CS\}$ for both extremes under a constraint of simulation budget. To efficiently resolve practical problems that have typically large stochastic noise, we developed a procedure utilizing the UE approach with high robustness against noise. We characterize the uncertainty of each design based on statistical hypothesis test and p -value, and present a sequential procedure that gradually allocates a few further replications according to the evaluated uncertainty. Various experimental results on both numerical benchmark and practical problems demonstrate the necessity of the proposed procedure compared to previous methods, such as the OCBA_{bw} procedure.

The remainder of this paper is organized as follows: Section II defines the problem, and Section III proposes the simulation budget allocation procedure. Section IV exhibits the experimental results. Finally, the conclusion is given in Section V.

II. PROBLEM DEFINITION

This paper uses the following basic notations:

- T : total number of simulation replications (i.e., the given simulation budget);
- k : number of simulation input designs;
- x_i : simulation input design, where subscript i represents the design index (i.e., $i \in \{1, \dots, k\}$);
- Θ : set of input designs, $\Theta = \{x_1, \dots, x_k\}$;
- Y_{ij} : simulation output sample of x_i in the j th replication, $Y_{ij} \sim \mathcal{N}(\mu_i, \sigma_i^2)$;
- μ_i : expected value of Y_{ij} (i.e., the mean of Y_{ij}), $\mu_i = E[Y_{ij}]$;
- σ_i^2 : variance of Y_{ij} , $\sigma_i^2 = \text{Var}[Y_{ij}]$;
- N_i : number of collected output samples of x_i (i.e., number of actually allocated replications at x_i);
- $\bar{\mu}_i$: sample average of N_i output samples of x_i , $\bar{\mu}_i = 1/N_i \cdot \sum_{j=1}^{N_i} Y_{ij} \sim \mathcal{N}(\mu_i, \sigma_i^2/N_i)$;
- s_i^2 : sample variance of N_i output samples of x_i , $s_i^2 = \sum_{j=1}^{N_i} (Y_{ij} - \bar{\mu}_i)^2 / (N_i - 1)$.

Based on common assumptions used in the R&S literature, a simulation output sample is supposed to follow independent and identically distributed normal distribution with unknown μ_i and σ_i^2 for every replication of the same x_i . That is, the simulation model has a Gaussian noise with zero mean in its output. This normality assumption can be reasonable in simulation fields as Y_{ij} is typically obtained as an average value or batch means, so the central limit theorem holds [8]. In addition, we assume that no prior knowledge of μ_i and σ_i^2 for each x_i is given before conducting simulations.

Suppose there is only one extreme design on each side, and denote that the design with the smallest μ_i is the best and the inverse (i.e., greatest μ_i) is the worst. Then, based on the obtained simulation results of the designs, the selected best design x_b and worst design x_w are defined as follows:

$$x_b = \arg \min_{x_i \in \Theta} \bar{\mu}_i \quad \text{and} \quad x_w = \arg \max_{x_i \in \Theta} \bar{\mu}_i, \quad (1)$$

For the selections of x_b and x_w to be correct, the unknown μ_b should be smaller than the other designs' μ_i , and the unknown μ_w should be larger than the other designs' μ_i . Thus, similar to [13], $P\{\text{CS}\}$ for both selected extreme designs can be defined as follows:

$$P\{\text{CS}\} = P\{\mu_b < \forall \mu_i < \mu_w, i \in \{1, \dots, k\}, i \neq b, w\}. \quad (2)$$

Simply, $P\{\text{CS}\}$ converges to 1 as T goes to infinity; however, a simulation budget is typically a constraint in practical situations. To improve the efficiency of simulation, our objective is to maximize $P\{\text{CS}\}$ under a limited number of simulation replications T by allocating T intelligently to each design, which is defined as follows:

$$\arg \max_{N_1, \dots, N_k} P\{\text{CS}\} \text{ such that } \sum_{i=1}^k N_i = T \text{ and } \forall N_i \geq 0. \quad (3)$$

Here, the constraint $\sum_{i=1}^k N_i = T$ implicitly supposes that the simulation cost of each replication across designs is similar. Otherwise, this can be roughly considered in (3) by reflecting the proportion of the increasing simulation cost for a replication of design in its variance [8].

III. PROPOSED PROCEDURE

A. UNCERTAINTY EVALUATION

For the selected best design x_b , the following $k - 1$ relationships between μ_b and μ_i depending on (2) should be verified for the correct selection of x_b :

$$(\mu_b < \mu_1) \wedge \dots \wedge (\mu_b < \mu_i) \wedge \dots \wedge (\mu_b < \mu_w) \quad \text{where } i \in \{1, \dots, k\} \text{ and } i \neq b. \quad (4)$$

On the other hand, for the selected worst design x_w , the following $k - 1$ relationships between μ_w and μ_i depending on (2) should be verified for the correct selection of x_w :

$$(\mu_1 < \mu_w) \wedge \dots \wedge (\mu_i < \mu_w) \wedge \dots \wedge (\mu_b < \mu_w) \quad \text{where } i \in \{1, \dots, k\} \text{ and } i \neq w. \quad (5)$$

For the remaining $k - 2$ designs, which are not selected as x_b and x_w , the following two relationships depending on (2) should be verified for the correct selection of each x_i :

$$(\mu_b < \mu_i) \wedge (\mu_i < \mu_w). \quad (6)$$

As in the other R&S procedures, our proposed procedure also allocates a few further replications based on the current simulation results, such as $\bar{\mu}_i$ and s_i^2 , to maximize $P\{\text{CS}\}$. If the relationships for the correct selection of design can be verified with the current simulation results, it is no longer necessary to collect simulation output samples for that design. Otherwise, if we cannot determine whether the selection for design is correct based on the current results, it is necessary to allocate further replications and collect more samples. The concept of uncertainty in the UE approach suggests an allocation criterion based on a statistical hypothesis test and p -value to accomplish this allocation strategy [15]. Utilizing the UE approach, our proposed procedure defines the uncertainty of each design and allocates further replications sequentially depending on the uncertainty to maximize $P\{\text{CS}\}$.

The statistical hypothesis test is a frequentist method to verify the single relationship between μ_i and μ_j based on simulation results. For example, when verifying $\mu_i < \mu_j$, this relationship is set to the alternative hypothesis H_A , and the opposite relationship is set to the null hypothesis H_0 , as follows:

$$H_0 : \mu_i \geq \mu_j, \quad H_A : \mu_i < \mu_j. \quad (7)$$

Then, the p -value of the test $p_{i,j}$ can be calculated as follows:

$$p_{i,j} = F_v [(\bar{\mu}_i - \bar{\mu}_j) / s_{i,j}] \quad \text{where} \\ s_{i,j} = \sqrt{s_i^2 / N_i + s_j^2 / N_j} \quad \text{and} \\ v = \left\lfloor \frac{s_i^4 / [s_i^4 / (N_i^3 - N_i^2) + s_j^4 / (N_j^3 - N_j^2)]}{s_{i,j}^4} \right\rfloor. \quad (8)$$

The function F_ν is the cumulative distribution function of the t -distribution with ν degrees of freedom. (If σ_i^2 and σ_j^2 are known, then the normal distribution can be used instead of the t -distribution.) Mathematically, $p_{i,j}$ is the probability of obtaining the current simulation results or more extreme results (which are less likely to be obtained) when assuming that H_0 is true. However, in the framework of the hypothesis test, it can be interpreted as the degree to which the current results can be considered significant evidence to verify $\mu_i < \mu_j$. The lower the value of $p_{i,j}$, the less likely it is that the current results will be obtained under the assumption that H_0 is true; however, the current results have been actually observed, so they can be significant evidence to verify $H_A: \mu_i < \mu_j$. On the other hand, as the value of $p_{i,j}$ increases, obtaining the current results becomes more probable under this assumption, so the obtained results cannot be significant evidence. That is, it is uncertain that H_A will be verified based on the current results (but, this does not mean that H_A is false). Although the possible range of $p_{i,j}$ is between 0 and 1, the meaningful range in the hypothesis test framework is between 0 and 0.5. This is because $1 - p_{i,j}$ is the same as $p_{j,i}$, which is the p -value for verifying the opposite of (7). Thus, $p_{i,j} = p_{j,i} = 0.5$ indicates the most uncertain case in which it is impossible to determine whether $\mu_i < \mu_j$ or $\mu_i > \mu_j$ based on the current simulation results [24].

If there is only one relationship (e.g., $\mu_i < \mu_j$) to be verified for the correct selection of design x_i , then $p_{i,j}$ can be the allocation criterion for the mentioned allocation strategy. That is, the higher value of $p_{i,j}$, the more insufficient the simulation results of x_i , so more output samples of x_i should be collected by allocating further replications. However, as shown in (4)-(6), several relationships should be verified simultaneously for the correct selection of each design in this problem. In this case, the uncertainty of each design defined as the combination of p -values for verifying each relationship can be the allocation criterion. For the selected best design x_b , each of the $k - 1$ relationships in (4) can be verified by the hypothesis tests, and the p -value from each test is denoted by $p_{b,i}$, where $i \neq b$. Each value of $p_{b,i}$ can be used to assess the degree to which the current simulation results of x_b can be significant evidence for each relationship in (4). If the results of x_b are not evidence for verifying even one relationship in (4), the results cannot be evidence for the correct selection of x_b , and the selection becomes uncertain; thus, the uncertainty of x_b , δ_b can be defined as the maximum value among $p_{b,i}$ s, as follows:

$$\delta_b = \max [p_{b,1}, \dots, p_{b,i}, \dots, p_{b,w}] \quad \text{where } i \neq b. \quad (9)$$

As the value of δ_b decreases to zero, the current simulation results of x_b can be significant evidence for the selection of x_b ; thus, it is no longer necessary to allocate further replications to x_b . However, as the value of δ_b increases to 0.5, at least one relationship in (4) is uncertain. Accordingly, we cannot determine whether the selection of x_b is correct based on the current results of x_b . Therefore, it is necessary to allocate further replications to x_b and update the insufficient results.

Similarly, for the selected worst design x_w , each of the $k - 1$ relationships in (5) can be verified by the hypothesis tests, and the p -value from each test is denoted by $p_{i,w}$, where $i \neq w$. In common with x_b , the current simulation results of x_w should be significant evidence for verifying every relationship in (5) for the correct selection of x_w , so the uncertainty of x_w , δ_w can be defined as the maximum value among $p_{i,w}$ s, as follows:

$$\delta_w = \max [p_{1,w}, \dots, p_{i,w}, \dots, p_{b,w}] \quad \text{where } i \neq w. \quad (10)$$

For each x_i ($i \neq b$ and $i \neq w$) of the remaining $k - 2$ designs that are not selected as x_b or x_w , the two relationships in (6) can be verified by hypothesis tests. p -values from both tests are denoted as $p_{b,i}$ and $p_{i,w}$. The current simulation results of x_i should be significant evidence verifying the two relationships in (6) for the correct selection of x_i . Therefore, the uncertainty of x_i , δ_i can be defined as the maximum value between $p_{b,i}$ and $p_{i,w}$, as follows:

$$\delta_i = \max [p_{b,i}, p_{i,w}] \quad \text{where } i \neq b \text{ and } i \neq w. \quad (11)$$

Strictly speaking, the uncertainty of each design in (9)-(11) includes the significance degree of not only this design's simulation results, but also another design's results. This is because the p -value in (8) is calculated using the two designs' results. However, if the estimated uncertainty is not very large or has specific extreme cases, it does not have a significant impact on allocation [25]. In addition, our objective is to develop a simple and practical means for allocating simulation budget to significantly improve the efficiency of simulation, as demonstrated in the experimental results later; thus, we neglect this problem in the proposed procedure.

B. ALLOCATION POLICY

To maximize $P\{\text{CS}\}$, the simulation results of all designs should be significant evidence verifying the relationships for the correct selection of each design shown in (4)-(6). To this end, more replications should be allocated to the designs of which the uncertainty has a relatively high value depending on the features of uncertainty. However, due to the type I error included in the uncertainty, a low uncertainty value also needs to be considered in the allocation. For example, suppose x_1 is the actual best design we must find, but unfortunately, after collecting some samples, the current simulation results of x_1 are very poor due to large stochastic noise. That is, x_1 is not selected as x_b , and $\bar{\mu}_1$ is far from both $\bar{\mu}_b$ and $\bar{\mu}_w$ (i.e., $\bar{\mu}_b \ll \bar{\mu}_1 \ll \bar{\mu}_w$), so the uncertainty of x_1 , δ_1 has low value according to (11). Although x_1 requires updating of its poor simulation results more than any other design, it cannot be allocated further replications due to its lower uncertainty. This type I error results in a reduction in $P\{\text{CS}\}$. Since the simulation results in type I error are typically very biased, just one or two additional samples are sufficient to check and eliminate the error. In other words, it is necessary to allocate a few additional replications to the designs of which the uncertainty has a relatively low value to increase $P\{\text{CS}\}$.

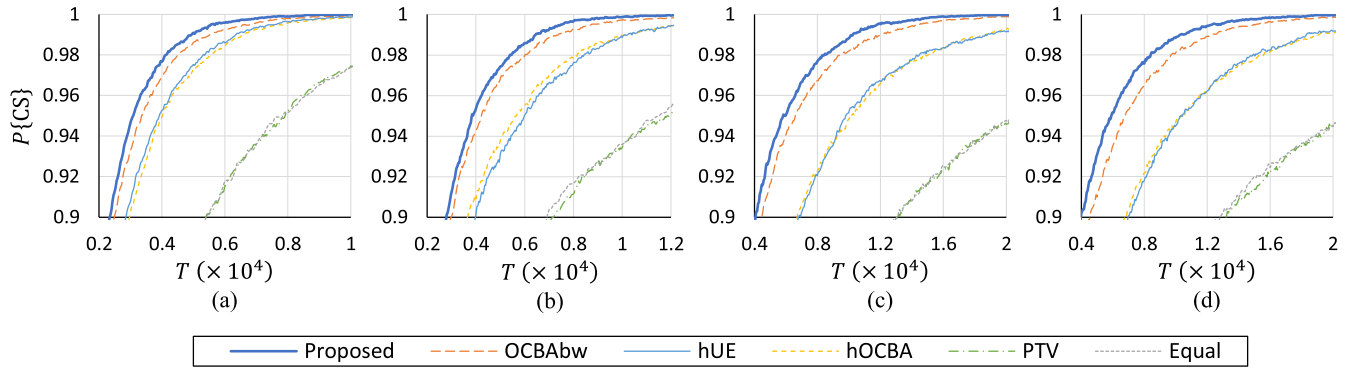


FIGURE 1. Graphs (a)-(d) illustrates the estimated value of $P\{CS\}$ versus T for the benchmark problems: (a) EV(L), (b) IV(L), (c) flat-case(L), and (d) steep-case(L).

Accordingly, a heuristic allocation policy to accomplish the mentioned allocation strategy based on uncertainty can be defined as follows [15]:

$$a_i/a_j = (\delta_i/\delta_j)^C \text{ where } i, j \in \{1, \dots, k\} \text{ and } \Delta = \sum_{i=1}^k a_i. \tag{12}$$

Given the number of further assignable replications Δ , a_i is the number of further allocated replications at x_i depending on its uncertainty δ_i . In the policy of (12), the coefficient C is a parameter that can adjust allocation depending on uncertainty, where the range of C is between zero and infinity. As the value of C increases, the difference of uncertainty is emphasized, so Δ is concentrated on a few designs that have relatively higher uncertainty. On the other hand, as the value of C decreases to zero, the difference of uncertainty is thinned, so the designs with relatively low uncertainty are more likely to be allocated a few further replications. The optimal value of C that maximizes $P\{CS\}$ is problem-specific. A high value can be effective in small-noise environments where a type I error is less likely to occur due to the relatively precise simulation results. However, in large noise environments where a type I error is more likely to occur due to inaccurate simulation results, a low value can be better.

In this paper, the default value of C is recommended as one. Although this value is not always optimal, it exhibits sufficiently good efficiency compared to the other R&S procedures, as shown in Table 2 and Fig. 1. In addition, if we consider selecting just a single extreme design such as the best one, the defined uncertainty in (9) and (11) becomes identical to that of the UE procedure finding the single best design. The UE procedure utilizes the allocation policy of (12) with $C = 1$ derived from the simulation-based optimizations; thus, this consistency gives a basis for this recommendation. In addition, the consistency can be a potential advantage of the proposed procedure over the OCBAbw procedure, which becomes identical to the OCBA procedure when considering a single extreme design. This is because the UE procedure has been proven empirically superior to the OCBA procedure in the presence of large noise [15].

C. SEQUENTIAL PROCEDURE

Algorithm 1 represents the proposed procedure for selecting both extreme designs simultaneously using the defined uncertainty of (9)-(11) and the allocation policy of (12). Prior to the iterative allocation of Δ , the proposed procedure collects n_0 output samples for each design to obtain minimal simulation results to evaluate the uncertainty. Then, until the given simulation budget T is exhausted, the procedure iterates 1) evaluating the uncertainty of each design, 2) allocating a few further replications of Δ depending on the policy, and 3) updating the simulation results. As the allocation repeats, the simulation results for each design are selectively and gradually updated to become significant evidence for their correct selections. Meanwhile, the first evaluations of δ_b and δ_w in the seventh line of Algorithm 1 decrease the computational cost

Algorithm 1 Select Both Extreme Designs Simultaneously
Control parameters: $T (\geq kn_0)$, n_0 , Δ , and C (default value is 1)

- Procedure:**
- 1: **simulate** n_0 times for each $x_i, i \in \{1, \dots, k\}$
 - 2: **update** $\bar{\mu}_i, s_i$, and N_i for $\forall i \in \{1, \dots, k\}$
 - 3: **select** x_b and x_w with (2)
 - 4: **while** $\sum_{i=1}^k N_i < T$ **do**
 - 5: **set** $\Delta \leftarrow \min(T - \sum_{i=1}^k N_i, \Delta)$
 - 6: **evaluate** δ_b and δ_w for x_b and x_w with (9) and (10)
 - 7: **evaluate** δ_i for $\forall i \in \{1, \dots, k\}, i \neq b$, and $i \neq w$ with (11)
 - 8: **calculate** a_i for $\forall i \in \{1, \dots, k\}$ with (12)
 - 9^a: **simulate** $\text{round}(a_i)$ times for each $x_i, i \in \{1, \dots, k\}$
 - 10: **update** $\bar{\mu}_i, s_i$, and N_i for $\forall i \in \{1, \dots, k\}$
 - 11: **select** x_b and x_w with (2)
 - 12: **end while**
 - 13: **return** x_b and x_w

^aDue to the robustness of the sequential procedure, any reasonable method for converting a real number a_i to an integer has not a significant impact on $P\{CS\}$. In Algorithm 1, we used the round function and scattered the rounding errors to several designs that have relatively larger a_i .

of the proposed procedure. This is because the uncertainty of the remaining $k - 2$ designs in the eighth line can be evaluated at no extra cost using $p_{b,i}$ and $p_{i,w}$ calculated in the previous line.

The sequential allocation procedure is controlled by two parameters: n_0 (the initial number of replications) and Δ (the number of further replications to be allocated per iteration). An optimal setting of n_0 and Δ is problem-specific. However, a recommended choice of n_0 is between 5 and 20 [26]. The recommended Δ value is between $0.5k$ and $0.05T$ [15] (see [15] for a detailed analysis of the UE approach according to n_0 and Δ). Meanwhile, a larger setting of n_0 and Δ than recommended might be effective in the presence of large stochastic noise. Increasing n_0 can decrease the possibility of type I errors by improving the initial simulation results. Increasing Δ raises the likelihood that one or two further replications will be allocated to designs with relatively low uncertainty, thereby checking and reducing the errors.

IV. EXPERIMENTS

In this section, we describe the experimental results on both numerical benchmark and practical problems that demonstrate the necessity of the proposed procedure.

A. BENCHMARK PROBLEMS

Table 1 presents the benchmark problems, which have been used to evaluate various R&S procedures, including the OCBA_{bw} procedure [8], [21]. All five problems have 10 designs with increasing μ_i according to the design index i , of which x_1 and x_{10} are both extreme designs that we have to select. The designs in the equal variance (EV) problem have the same variance, whereas in the increasing variance (IV) and decreasing variance (DV) they have different variances that increase or decrease linearly according to i . While μ_i in these three problems increases linearly, μ_i in the flat- and steep-case problems increases concavely and convexly, respectively. Each problem has a large noise version denoted by L, which has relatively larger variance than its original version. Thus, the experimental results for the L version may well indicate the effectiveness of the proposed procedure.

TABLE 1. Five numerical benchmark problems.

Problem	k	μ_i	σ_i	$\sigma_i(L)$
Equal variance (EV)	10	i	6	10
Increasing variance (IV)	10	i	i	$5 + i$
Decreasing variance (DV)	10	i	$11 - i$	$16 - i$
Flat-case	10	$10 - 3\sqrt{10 - i}$	6	10
Steep-case	10	$1 + 3\sqrt{i - 1}$	6	10

For these five benchmarks, we compared the efficiency of the proposed procedure with other procedures: OCBA_{bw}, hOCBA, hUE, PTV, and the equal allocation procedures. The hOCBA and hUE procedures are variants of the OCBA and UE procedures that find only the best design. They use the OCBA or UE procedures twice to find both extreme designs.

That is, in each iteration of the sequential procedure, half of the given Δ is used to find the best design, and the other half is used to find the worst one. Meanwhile, since the OCBA_{bw} procedure is reduced to the OCBA_{bw} procedure when $m = n = 1$, it was not considered in this experiment. For a fair comparison, n_0 and Δ in every procedure were set to 10 and 20 according to the typical settings of the five problems [8]. We estimated $P\{CS\}$ for each procedure with 10,000 independent repeated experiments while varying T . The several results for the L version of the problems are shown in Fig. 1. In addition, for numerical comparisons, the simulation budget required to find both extreme designs correctly (i.e., to achieve $P\{CS\}$ of 0.99) is measured in Table 2. Our experimental results indicate the superior efficiency of the proposed procedure over other procedures.

TABLE 2. Simulation budget T required to achieve $P\{CS\}$ of 0.99 for benchmark problems.

Problems		Proposed	OCBA _{bw}	hUE	hOCBA	PTV	Equal
EV	T req.	1,920	2,040	2,360	2,500	4,800	4,900
	Ratio ^a	-	1.06	1.23	1.30	2.50	2.55
EV(L)	T req.	4,940	5,560	6,340	6,600	13,720	13,800
	Ratio	-	1.13	1.28	1.34	2.78	2.79
IV	T req.	2,360	2,480	4,389	4,360	9,760	10,300
	Ratio	-	1.05	1.86	1.85	4.14	4.36
IV(L)	T req.	6,440	7,400	10,140	10,160	23,600	23,000
	Ratio	-	1.15	1.57	1.58	3.66	3.57
DV	T req.	2,300	2,440	4,300	4,280	9,680	9,650
	Ratio	-	1.06	1.87	1.86	4.21	4.2
DV(L)	T req.	6,640	7,340	10,340	10,440	22,920	23,100
	Ratio	-	1.11	1.56	1.57	3.45	3.48
flat-case	T req.	4,060	4,440	6,640	6,900	14,540	14,500
	Ratio	-	1.09	1.64	1.7	3.58	3.57
flat-case(L)	T req.	10,260	12,040	18,680	18,700	43,060	41,900
	Ratio	-	1.17	1.82	1.82	4.20	4.08
steep-case	T req.	3,900	4,540	6,760	7,020	15,640	14,000
	Ratio	-	1.16	1.73	1.80	4.01	3.59
steep-case(L)	T req.	10,280	12,400	18,320	19,100	45,000	41,500
	Ratio	-	1.21	1.78	1.86	4.38	4.04

^aThis denotes the ratio of T required in the proposed procedure and that required in the other procedure (i.e., $T_{\text{other}}/T_{\text{proposed}}$).

The results of the hUE and hOCBA procedures represent why a procedure for selecting both extreme designs simultaneously is necessary. Except for the EV problem, the simulation budget for selecting the best and worst designs is not the same due to the inconsistent μ_i spacing (the flat- and steep-case problems) or the different variance (the IV and DV problems), which is a common case in practice. As shown in Table 2, compared to the results of the EV problem, the proposed procedure is more efficient than both procedures in these practical cases. For example, in the EV problem, the hUE and hOCBA procedures consumed 1.23 times and 1.30 times more replications than the proposed procedure, respectively; however, in the IV problem, these numbers increased significantly to 1.86 and 1.87, respectively.

In addition, the results of the OCBA_{bw} procedure, especially in the L version of the problems, emphasize the high robustness against noise of the proposed procedure.

For example, in the flat case problem, while the OCBA_{bw} procedure consumed 1.09 times more replications than the proposed procedure, the number increased to 1.17 in the ‘(L)’ version, as shown in Table 2. Similar trends were observed in the other problems. Compared to the OCBA_{bw} procedure, the higher robustness of the proposed procedure is attributed to considering additional information when allocating further replications. Although both procedures take into account the sample mean $\bar{\mu}_i$ and the sample variance s_i^2 , the proposed procedure considers the number of collected samples N_i as well, as shown in (8). However, the OCBA_{bw} procedure uses only $\bar{\mu}_i$ and s_i^2 as specified in its main allocation rule below [21]:

$$\frac{N_i^*}{N_j^*} = \frac{s_i^2 / \min((\bar{\mu}_b - \bar{\mu}_i)^2, (\bar{\mu}_i - \bar{\mu}_w)^2)}{s_j^2 / \min((\bar{\mu}_b - \bar{\mu}_j)^2, (\bar{\mu}_j - \bar{\mu}_w)^2)}. \quad (13)$$

Here, N_i^* is the optimal number of replications that should be assigned to x_i to maximize a lower bound of $P\{CS\}$ based on the current simulation results. That is, N_i^* is different from N_i , the number of actually allocated replications so far, and the further replications actually allocated to x_i is defined as $\max(0, N_i^* - N_i)$ in the OCBA_{bw} procedure. Simply put, while the proposed procedure takes into account the precision of the sample mean $s_i / \sqrt{N_i}$ when allocating Δ , the OCBA_{bw} procedure cannot consider it (s_i^2 itself is not enough to indicate the precision of $\bar{\mu}_i$). Accordingly, the OCBA_{bw} procedure is likely to make an inefficient allocation due to an inaccurate $\bar{\mu}_i$ value, which often occurs in the presence of large stochastic noise.

For example, in the EV problem, suppose the observed $\bar{\mu}_1$ of the true best design x_1 is unfortunately high (i.e., close to 5). Then, the OCBA_{bw} procedure cannot allocate further replications to x_1 due to the inaccurate value of $\bar{\mu}_1$. Since $\bar{\mu}_1$ cannot be updated, the situation in which no further replications are allocated to x_1 continues. If T is infinite, x_1 will be allocated further replications depending on changes in the simulation results of the other designs and will be selected as x_b eventually. However, numerous replications are wasted on other marginal designs in the meantime. On the other hand, our proposed procedure, which considers the precision of $\bar{\mu}_1$, can allocate further replications quickly to x_1 because the precision of $\bar{\mu}_1$ that is not updated becomes relatively lower. Thereby, it can select x_1 as x_b efficiently. Meanwhile, if the noise level is small, considering N_i may reduce efficiency. However, this reduction is insignificant since the required T for the correct selection is small in this case. In addition, most complex practical problems have large stochastic noise. Thus, the proposed procedure can be effective, as shown in the next subsection.

B. PRACTICAL PROBLEM: MUNICIPAL WASTE MANAGEMENT IN A CITY

As urbanization intensifies in modern society, municipal waste management in dense residential environments has become a social problem. As it is directly related to the

satisfaction and hygiene of residents, many complaints are generated in this regard, which wastes administrative power. The government allocates garbage trucks depending on the number of residents and periodically collects municipal waste; however, various variables, such as regional characteristics, residential patterns, local developments, etc., make it difficult to manage effectively. A studio apartment complex in South Korea is a representative place where many complaints against municipal waste management arise because inexpensive rent often changes the residential patterns. To reduce complaints, discrete-event simulations can be effectively used for preliminary analysis of a given management policy [27]. Given 30 scenario designs that reflect the possible residential patterns, practitioners are interested in not only the best design, but also the worst design under which the number of complaints is maximized. That is because even if the number of complaints in the best-case scenario is very small, the evaluation of this policy should be reconsidered if its worst-case scenario frequently occurs. Our proposed procedure can be effectively applied to such an analysis of municipal waste management policy based on the simulation.

As shown in Fig. 2, the relatively large σ_i compared to the small difference in μ_i between both extreme designs and their neighborhood designs acts as a large stochastic noise making it difficult to find the extreme designs. To demonstrate the necessity of the proposed procedure for such a practical problem, other procedures used in the benchmark experiments were applied together, and Fig. 3 represents the results. While the proposed procedure required only 5,880 replications to correctly find the best and worst designs (i.e., to achieve $P\{CS\}$ of 0.99), the OCBA_{bw} procedure required 8,760, and the hUE and hOCBA procedures required 7,890 and 12,450, respectively. With no R&S procedures applied, 50,700 replications were required. This simulation model is expensive per simulation run (i.e., about 1 min.) due to the interactions between many agents; thus, applying the proposed procedure can greatly reduce the time needed to analyze the municipal waste management policy by enhancing the simulation’s efficiency.

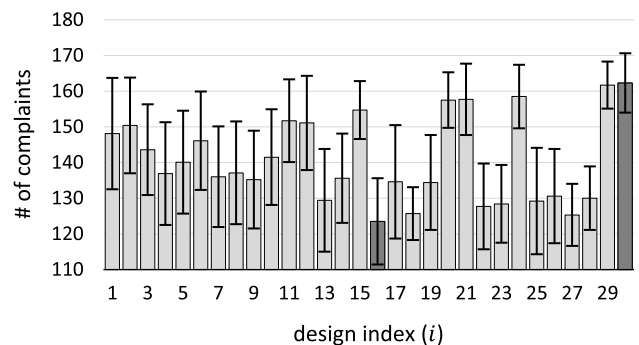


FIGURE 2. The precisely estimated value of μ_i (i.e., the number of complaints) and σ_i for 30 given designs using many repeated simulations [27]. The two dark gray bars represent the results of both extreme designs (i.e., x_{16} and x_{30}) that we have to select.

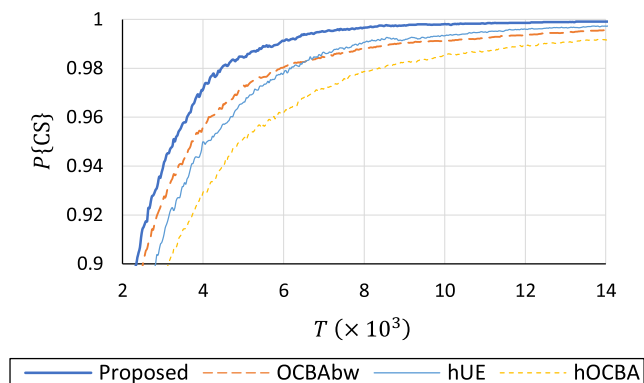


FIGURE 3. The estimated value of $P\{CS\}$ versus T for the municipal waste management problems in Fig. 2. The $P\{CS\}$ was estimated over 1,000 independent repeated experiments.

Like this practical problem, the proposed procedure is effective in problems where the simulation model has large stochastic noise and high evaluation costs. A digital twin, drawing attention in the era of the fourth industrial revolution, is an example that meets these conditions. The digital twin is a simulation model in cyberspace corresponding to a system in physical space. It is continuously updated to maintain synchronization with the real system, and the best and worst control settings should be found quickly based on the updated model to optimize the system. In this case, the superior efficiency of the proposed procedure can help the use of digital twins, especially for systems that demand high reliability.

V. CONCLUSION

Finding both extreme input designs has tremendous potential in discrete-event simulation as well as in other areas. In this paper, we proposed an efficient R&S procedure for selecting the best and worst designs simultaneously from a finite set of design alternatives under large stochastic noise. To maximize $P\{CS\}$ in the constraint of simulation budget, we described evidence of correct selection using uncertainty based on the statistical hypothesis test and p -values. Then, we proposed a procedure that sequentially allocates a few further simulation replications depending on the evaluated uncertainty, such that the simulation results can provide significant evidence for the correct selection. The proposed procedure is superior to using basic R&S procedures (e.g., OCBA or UE procedures) twice because it efficiently distributes the budget to both extreme designs. Furthermore, since it additionally considers the number of allocated replications so far in allocation, it is more efficient than the OCBAbw procedure in the presence of large stochastic noise. The experimental results on benchmark problems demonstrate this, and the municipal waste management problem emphasizes the necessity of the proposed procedure for practical simulation problems. As mentioned earlier, discrete-event simulation is used in a variety of engineering fields, and the philosophy of selecting both extreme designs simultaneously also has tremendous

potential for application. As a future study, we will apply and customize the proposed procedure for improving the efficiency of learning automata [28], network resource allocation strategies [29], and digital twins.

ACKNOWLEDGMENT

The authors would like to thank the Editor, an Associate Editor, and the anonymous reviewers for their valuable comments and constructive suggestions.

REFERENCES

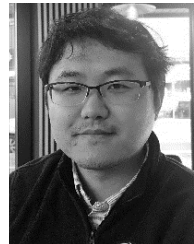
- [1] Y. Liu, K. Wang, L. Ge, L. Ye, and J. Cheng, "Adaptive evaluation of virtual machine placement and migration scheduling algorithms using stochastic Petri nets," *IEEE Access*, vol. 7, pp. 79810–79824, 2019.
- [2] B. G. Kang, K.-M. Seo, and T. G. Kim, "Model-based design of defense cyber-physical systems to analyze mission effectiveness and network performance," *IEEE Access*, vol. 7, pp. 42063–42080, 2019.
- [3] S. H. Mousavi, M. Davari, and H. J. Marquez, "An innovative event-based filtering scheme using H_∞ performance for stochastic LTI systems considering a practical application in smart modernized microgrids," *IEEE Access*, vol. 7, pp. 48138–48150, 2019.
- [4] L. Wang, R. Iida, and A. M. Wyglinski, "Vehicular network simulation environment via discrete event system modeling," *IEEE Access*, vol. 7, pp. 87246–87264, 2019.
- [5] S. Yang, K. Song, and G. Zhu, "Stochastic process and simulation of traction load for high speed railways," *IEEE Access*, vol. 7, pp. 76049–76060, 2019.
- [6] H. Xiao, L. H. Lee, and K. M. Ng, "Optimal computing budget allocation for complete ranking," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 2, pp. 516–524, Apr. 2014.
- [7] H. Xiao, H. Chen, and L. H. Lee, "An efficient simulation procedure for ranking the top simulated designs in the presence of stochastic constraints," *Automatica*, vol. 103, pp. 106–115, May 2019.
- [8] C.-H. Chen and L. H. Lee, *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*, vol. 1. Singapore: World Scientific, 2011.
- [9] J. Xu, E. Huang, C.-H. Chen, and L. H. Lee, "Simulation optimization: A review and exploration in the new era of cloud computing and big data," *Asia-Pacific J. Oper. Res.*, vol. 32, no. 3, Jun. 2015, Art. no. 1550019.
- [10] S. Gao, L. H. Lee, C.-H. Chen, and L. Shi, "A sequential budget allocation framework for simulation optimization," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 2, pp. 1185–1194, Apr. 2017.
- [11] J. Branke, S. E. Chick, and C. Schmidt, "Selecting a selection procedure," *Manage. Sci.*, vol. 53, no. 12, pp. 1916–1932, Dec. 2007.
- [12] S.-H. Kim and B. L. Nelson, "A fully sequential procedure for indifference-zone selection in simulation," *ACM Trans. Model. Comput. Simul.*, vol. 11, no. 3, pp. 251–273, Jul. 2001.
- [13] C.-H. Chen, J. Lin, E. Yücesan, and S. E. Chick, "Simulation budget allocation for further enhancing the efficiency of ordinal optimization," *Discrete Event Dyn. Syst.*, vol. 10, no. 3, pp. 251–270, Jul. 2000.
- [14] S. H. Choi and T. G. Kim, "Enhancing the noise robustness of the optimal computing budget allocation approach," *IEEE Access*, vol. 8, pp. 25749–25763, Jan. 2020.
- [15] S. H. Choi and T. G. Kim, "Efficient ranking and selection for stochastic simulation model based on hypothesis test," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 9, pp. 1555–1565, Sep. 2018.
- [16] S. H. Choi, K.-M. Seo, and T. G. Kim, "Accelerated simulation of discrete event dynamic systems via a multi-fidelity modeling framework," *Appl. Sci.*, vol. 7, no. 10, p. 1056, Oct. 2017.
- [17] W. G. Lou, "Study on layout planning of municipal sewage treatment plants by using osculating method," *Shanghai Environ. Sci.*, vol. 21, no. 6, pp. 327–329, 2002.
- [18] T. N. Flynn, J. J. Louviere, T. J. Peters, and J. Coast, "Best–worst scaling: What it can do for health care research and how to do it," *J. Health Econ.*, vol. 26, no. 1, pp. 171–189, Jan. 2007.
- [19] S. Zhang, P. Chen, L. H. Lee, C. E. Peng, and C.-H. Chen, "Simulation optimization using the particle swarm optimization with optimal computing budget allocation," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2011, pp. 4298–4309.

- [20] D. E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," *Found. Genetic Algorithms*, vol. 1, no. 1, pp. 69–93, 1991.
- [21] J. Zhang, L. Zhang, C. Wang, and M. Zhou, "Approximately optimal computing budget allocation for selection of the best and worst designs," *IEEE Trans. Autom. Control*, vol. 62, no. 7, pp. 3249–3261, Jul. 2017.
- [22] H. Xiao, S. Gao, and L. H. Lee, "Simulation budget allocation for simultaneously selecting the best and worst subsets," *Automatica*, vol. 84, pp. 117–127, Oct. 2017.
- [23] J. Zhang, Z. Li, C. Wang, D. Zang, and M. Zhou, "Approximate simulation budget allocation for subset ranking," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 1, pp. 358–365, Jan. 2017.
- [24] A. Hayter, *Probability and Statistics for Engineer and Scientists*, 4th ed., Boston, MA, USA: Cengage Learning, 2012.
- [25] S. H. Choi, C. Choi, and T. G. Kim, "An improved budget allocation procedure for selecting the best-simulated design in the presence of large stochastic noise," *IEEE Access*, vol. 7, pp. 154435–154446, 2019.
- [26] R. E. Bechhofer, T. J. Santner, and D. M. Goldsman, *Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons*. New York, NY, USA: Wiley, 1995.
- [27] C.-H. Lyoo, J.-H. Jung, C.-B. Choi, and E.-Y. Kim, "Modeling and simulation of municipal solid waste management system based on discrete event system specification," in *Proc. Symp. Simulation Architecture, Urban Design (SimAUD)*, May 2020.
- [28] J. Zhang, C. Wang, D. Zang, and M. Zhou, "Incorporation of optimal computing budget allocation for ordinal optimization into learning automata," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 2, pp. 1008–1017, Apr. 2016.
- [29] S. Kumar, R. Devaraj, and A. Sarkar, "A hybrid offline-online approach to adaptive downlink resource allocation over LTE," *IEEE/CAA J. Automatica Sinica*, vol. 6, no. 3, pp. 766–777, May 2019.



SEON HAN CHOI (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2012, 2014, and 2018, respectively.

He was a Teaching Assistant of computer programming with the School of Electrical Engineering, from 2013 to 2017. In 2018, he was a Postdoctoral Researcher with the Information and Electronics Research Institute, KAIST. From 2018 to 2019, he was a Senior Researcher with Korea Institute of Industrial Technology. In 2019, he joined the Department of IT Convergence and Application Engineering, Pukyong National University, Busan, South Korea, as an Assistant Professor. His current research interests include modeling and simulation of discrete-event systems, efficient simulation optimization under stochastic noise, evolutionary computing, and machine learning.



CHANGBEOM CHOI (Member, IEEE) received the B.S. degree in computer engineering from Kyung Hee University, Suwon, South Korea, in 2005, and the M.S. degree in computer science and the Ph.D. degree in electrical engineering from KAIST, Daejeon, South Korea, in 2007 and 2014, respectively.

He has been an Assistant Professor with Handong Global University, Pohang, South Korea, since 2014. His research interests include discrete-event systems modeling and simulation, experiment management, verification, validation, and accreditation, and agent-based simulation in various domains.

• • •