Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

UPPER-CONFIDENCE-BOUND PROCEDURE FOR ROBUST SELECTION OF THE BEST

Yuchen Wan
L. Jeff Hong

School of Data Science
Fudan University
440 Handan Road
Yangpu, Shanghai 200433, CHINA

Weiwei Fan

Advanced Institute of Business
Tongji University
1239 Siping Road,
Yangpu, Shanghai 200092, CHINA

ABSTRACT

Robust selection of the best (RSB) is an important problem in the simulation area, when there exists input uncertainty in the underlying simulation model. RSB models this input uncertainty by a discrete ambiguity set and then proposes a two-layer framework under which the best alternative is defined to have the best worst-case mean performance over the ambiguity set. In this paper, we adopt a fixed-budget framework to address the RSB problem. Specifically, in contrast with existing procedures, we develop a new robust upper-confidence-bound (UCB) procedure, named as R-UCB. We can show that, the R-UCB procedure successfully inherits the simplicity and convergence guarantee of the traditional UCB procedure. Furthermore, simulation experiments demonstrate that the R-UCB procedure numerically outperforms the existing RSB procedures.

1 INTRODUCTION

Ranking and selection (R&S) seeks to select the best alternative from a finite number of alternatives through repeatedly sampling a simulation model. The best alternative is defined to have the smallest mean performance. Conventionally, the simulation model as well as its input distributions is assumed to be known. However, when the input distributions are estimated from real-world data, we may suffer from the model ambiguity, due to the lack of data or the measurement error. This phenomenon is known as input uncertainty in the literature. The R&S problem concerning input uncertainty has drawn a lot of attention over the past few years. Fan et al. (2013) propose a robust selection-of-the-best (RSB) framework which models the input uncertainty by a discrete ambiguity set and selects the best alternative with the smallest worst-case mean performance over the ambiguity set. Particularly, this minimax RSB framework involves two layers, where the inner layer first identifies the worst-case input distribution for each alternative and then the outer layer compares the worst-case mean performances of alternatives to select the best one. In addition to R&S under input uncertainty, the RSB framework can be also implemented in other practical applications. For example, Lesnevski et al. (2007) take the robust perspective to select the portfolio with the minimum coherent risk measure which represents the maximum expected loss.

To address the RSB problem, a series of procedures have been proposed. Following the convention of R&S literature (Gabillon et al. 2012; Hunter and Nelson 2017; Hong et al. 2021), we classify existing RSB procedures into fixed-precision and fixed-budget procedures. Fixed-precision RSB procedures, e.g., the two-stage and sequential RSB procedures in Fan et al. (2020), attempt to achieve a pre-specified probability of correct selection (PCS) using as few simulation budget as possible. Meanwhile, fixed-budget R&S procedures attempt to optimize the quality of the final selection when the simulation budget is limited. Representative procedures include Gao et al. (2017) and Zhang and Ding (2016). These two procedures carefully extend the traditional optimal computing budget allocation (OCBA) and knowledge gradient (KG)

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

procedures to adapt themselves to the RSB framework. These extensions inspire us that existing procedures may be modified and applied to solve the RSB problem. In this paper, our goal is to design a new RSB procedure under the fixed-budget formulation.

More precisely, we design the new RSB procedure on basis of the upper-confidence-bound (UCB) procedure (Auer et al. 2002) from the field of multi-armed bandit (MAB), and the reason is two-folded. Firstly, the UCB procedure is simple to implement and promises to achieve a theoretical convergence. It works by providing a confidence interval for every arm to represent the uncertainty of the estimated means and choosing the arm with the highest upper confidence bound to achieve a balance between exploration and exploitation. No strict assumptions are required on the underlying simulation model (e.g., Gaussian distribution or bounded domain of simulation samples) and the calculation of confidence interval is often easy. Secondly but more importantly, there is a close link between R&S and MAB problems. Both of them stem from Bechhofer (1954) and aim to sequentially allocate the limited simulation budget among alternatives (or arms) to the most efficient extent. On account of the two reasons above, we think that it is interesting and promising to exploit a new RSB procedure based on the UCB procedure.

To accomplish this task, we need to overcome at least two challenges. Firstly, the outer and inner layers of RSB have different intrinsic goals, and henceforth different procedures are required for the two layers. To make it clear, in the inner layer, what we really care about is the worst-case mean performance of each alternative rather than the corresponding worst-case input distribution, because the worst-case mean is then used to guide the outer-layer selection. Meanwhile, the outer-layer problem refers to a classic R&S problem. Secondly, despite the similarity between R&S and MAB problems, they pursue different targets. The MAB assumes a regret at each stage whenever the best arm is not pulled, and therefore its target is to minimize the cumulative regret collected till the terminal stage. By contrast, the R&S assumes no intermediate regret along the dynamic sampling process, and the target is to minimize the simple regret at the terminal stage if the best alternative is incorrectly selected. Therefore, the key issue is to separately modify the traditional UCB procedure to solve the problems in the two layers of the RSB.

Notice that the traditional UCB procedure specifically sets the exploration rate to be a logarithm function, ensuring that the expected cumulative regret grows in the optimal order. As mentioned above, the two layers of RSB have different goals from the MAB problem. In our viewpoint, the most straightforward way is to properly alter the exploration rate in the traditional UCB procedure to make it suitable for solving the two-layer RSB problem. Driven by this insight, a new UCB procedure is proposed.

The rest of the paper is organized as follows. Section 2 formulates the RSB problem and constructs the connection between RSB and MAB problems. Section 3 presents the R-UCB procedure and proves its consistency. Section 4 provides the numerical experiments. Section 5 concludes the paper.

2 PROBLEM STATEMENT

2.1 Fixed-Budget Robust Selection of the Best

Suppose there is a finite set of alternatives $S = \{s_1, s_2, \ldots, s_k\}$. Let $g(s,\zeta)$ denote the performance of each alternative $s \in S$, where ζ is the input parameter following a probability distribution P_0. Assume that the best alternative is defined to have the smallest mean performance, i.e.,

$$\min_{s\in S} E_{P_0}[g(s,\zeta)].$$

Typically, $g(s,\zeta)$ is a black-box function and can only be observed via running simulation experiments. For instance, in a queueing system, s refers to the number of staff, ζ refers to the service time and $g(s,\zeta)$ refers to the waiting time. To drive the simulation, we often need to estimate the input distribution P_0 from the real-world data a priori. Practically, there might be ambiguity in the specification of the true input distribution P_0. In other words, the phenomenon of input uncertainty arises. Following the work of Fan et al. (2020), we model the input uncertainty by an ambiguity set $\mathscr{P} = \{P_1, \ldots, P_m\}$, which consists of m possible probability scenarios of P_0. Accordingly, the best alternative is re-defined as the alternative with the smallest worst-case mean performance over the ambiguity set, i.e.,

$$\min_{s\in S}\max_{P\in\mathscr{P}} E_P[g(s,\zeta)]. \tag{1}$$

This framework is known as the robust selection of the best (RSB). The RSB framework essentially contains two layers. The inner layer intends to find the worst-case probability scenario among the m possible probability scenarios in $\mathscr{P}$ for each alternative $s_i \in S$, and then the outer layer compares the worst-case probability scenarios of all k alternatives to select the best alternative.

For convenience of presentation, we denote by system (i,j) the pair of alternative $s_i \in S$ and probability scenario $P_j \in \mathscr{P}$, and hence there are km systems in total. From each system (i,j), we can collect independent and identically distributed (i.i.d.) samples $X_{ij,1}, X_{ij,2}, \ldots, X_{ij,l}, \ldots$, with mean $\mu_{ij} = E_{P_j}[g(s_i,\zeta)] = E[X_{ij,l}]$ and variance $\sigma_{ij}^2 = \mathrm{Var}_{P_j}[g(s_i,\zeta)] = \mathrm{Var}[X_{ij,l}]$. For the simplicity of notation, we assume that, for each alternative $i \in S$, its associated means are sorted in a descending order, i.e., $\mu_{i1} \geq \mu_{i2} \geq \ldots \geq \mu_{im}$, and the worst-case means of alternatives are sorted in an ascending order,i.e, $\mu_{k1} \geq \ldots \geq \mu_{21} > \mu_{11}$. Then, by Equation (1), alternative 1 is the unique best alternative and probability scenario 1 is its worst-case distribution. Therefore, system (1,1) is the best system among all km systems.

To address the RSB problem, we adopt a fixed-budget formulation and let N be the total simulation budget. Suppose that the budget N is sequentially allocated among systems over multiple stages, each of which is endowed with only one sample. Particularly, at each stage $t \leq N$, a budget allocation decision is made to tell which system should receive the next-stage sample based on all the past sample information. For each system (i,j), we use $n_{ij,t}$ to denote the number of samples that have been allocated to system (i,j) up to stage t and let $\bar{X}_{ij}(n_{ij,t})$ denote the corresponding sample mean. Then, at the terminal stage N when the total budget is exhausted, one system, denoted by $(\hat{i}^*, \hat{j}^*)$, is recommended as the best system based on the sample means of all systems, i.e., $\bar{X}_{ij}(n_{ij,N})$ for all $i = 1,2,\ldots,k$, $j = 1,2,\ldots,m$. Back to the RSB framework, we correspondingly select alternative $\hat{i}^*$ as the best alternative.

Apparently, the selected alternative $\hat{i}^*$ may not be the true best due the noise of the simulation samples. We measure the quality of the final selection by the expected opportunity cost (EOC), which refers to the expected difference of the worst-case means between the selected alternative $\hat{i}^*$ and the true best alternative 1, i.e.,

$$\mathrm{EOC} = E[\mu_{\hat{i}^*,1} - \mu_{11}].$$

Our goal is to design an optimal or near-optimal budget allocation procedure that can minimize the EOC of the final selection using the limited budget N.

2.2 Connection between R&S and MAB

By equation (1), the RSB problem involves two R&S problems due to its two-layer structure. As a building block, we may separetely consider the R&S problem in each layer. As stated in Section 1, one particular MAB procedure, i.e., the UCB procedure, could be a potentially powerful procedure to solve the R&S problem. Before addressing this issue in detail, we would like to first illustrate the connection between R&S and MAB problems.

The focus of both R&S and MAB problems is on designing a proper allocation strategy that can make the best use of the limited simulation budget N. What differentiates them from each other is the way to measure the performance of a given allocation strategy. Notice that, at each stage $t \leq N$, if the true best system is not recommended, a regret $r(t)$ which is defined by the mean gap between the best system and the recommended system would occur. The MAB problems measure an allocation strategy by the cumulative regret over all the N stages, namely, $R(N) = \sum_{t=1}^{N} r(t)$. By contrast, R&S problems measure the final selection by the (simple) single-stage regret $r(N)$ at the terminal stage N. Regardless of the different measurements used, we see a strong connection between them from the following Lemma.

Lemma 1 (Bubeck et al. 2009) For all functions $\varepsilon : \{1,2,\ldots\} \to \mathbb{R}$, there exist constants $C, D, \Delta > 0$ such that

$$E[R(n)] \leqslant C\varepsilon(n),\, E[r(n)] \geqslant \tfrac{\Delta}{2}e^{-D\varepsilon(n)}.$$

Lemma 1 reveals an interesting fact on the cumulative regret and the simple regret. The smaller the expected cumulative regret is, the larger the expected simple regret is. As traditional MAB procedures are typically designed to minimize the expected cumulative regret, they might result in an undesired large expected simple regret when the MAB problems are directly applied to address the R&S problem. Therefore, proper modifications are needed to translate the MAB procedures to solve the R&S problem.

3 ROBUST UCB PROCEDURE

In this section, our main task is to propose a robust UCB (R-UCB) procedure for the RSB problem. Due to the two-layer structure of the RSB problem, the R-UCB procedure needs to make two allocation decisions at each stage $t \leq N$ of the sequential sampling process. Firstly, in the inner layer, the procedure decides the probability scenario to sample from for each alternative $s_i \in S$. Secondly, in the outer layer, the procedure compares the probability scenarios of all k alternatives returned in the inner layer to produce the next-stage sampling decision. Combining these two allocation decisions, the procedure ultimately determines the alternative as its associated probability scenario which we will next simulate from.

3.1 Outer Layer: Selection and Minimization

In this section, we start by separately considering the two allocation decisions made in the two layers. Suppose for now that, at each stage t and for each alternative $s_i \in S$, a probability scenario $P_{j_t^i}$ or equivalently system (i, j_t^i) has been recommended in the inner layer. Then, the outer layer compares the k systems, i.e., $(1, j_t^1), (2, j_t^2), \dots, (k, j_t^k)$, to determine which system will be allocated the next sample. Obviously, the outer-layer problem performs exactly as a traditional R&S problem.

In solving the outer-layer minimization problem, the UCB procedure chooses to sample the system with the smallest lower confidence bound of means, i.e.,

$$(i_t, j_t) = \arg \min_{(i, j_t^i): i=1,2\dots,k} \left(\bar{X}_{i,j_t^i}(n_{ij_t^i,t}) - \sigma_{i,j_t^i}\sqrt{2\frac{a_t}{n_{i\cdot,t}}} \right),$$

for each alternative $i = 1, 2, \dots, k$. Here $n_{i\cdot,t} = \sum_{j=1}^{m} n_{ij,t}$ denotes the total simulation budget of all systems related to alternative i up to stage t and a_t refers to the exploration rate of the procedure.

Recall that the traditional UCB procedure is designed to minimize the expected cumulative regret in MAB, as introduced in Section 2.2. It is well known in the literature that the expected cumulative regret grows in the order no smaller than $\log t$ (Auer et al. 2002). To achieve this optimal order, the traditional UCB procedure typically sets the exploration rate $a_t = \log t$. However, such an exploration rate is not suitable for the outer-layer R&S problem whose objective is to minimize the expected simple regret. One natural way to translate the UCB procedure is by carefully adjusting the exploration rate a_t in the traditional UCB. Combined with Lemma 1, it can be understood that, a larger exploration rate leads to a larger expected cumulative regret and consequentially a smaller expected simple regret. In light of this, we set the exploration rate a_t in the order of t.

3.2 Inner Layer: Estimation and Maximization

Notice that, in the RSB problem, our goal is to select the best alternative rather than its corresponding worst-case probability scenario. Therefore, the inner-layer problem is more like a maximum-estimation problem rather than a R&S problem. Similar to Section 3.1, we may properly adjust the exploration rate in the UCB procedure to make it suitable for maximum estimation. Recently, Liu et al. (2019) have discussed this issue but their goal is to estimate the coherent risk measure of a portfolio. The coherent risk measure (Artzner et al. 1999), as a measurement of financial risk, represents the expected maximum loss of a portfolio and it is similar to our inner-layer problem. In their paper, they find that the proper exploration rate for maximum estimation should take value in the order ranging from $\log t$ to $t^{1/2-\delta}$ with $0 < \delta < 1/2$. Driven by their result, we set our inner-layer exploration rate for each alternative s_i as $\log n_{i.}$, where $n_{i.}$ denotes the total budget allocated to all systems of alternative s_i.

3.3 The Procedure

According to the discussion above, we are ready to present our R-UCB procedure as follows. This procedure starts with an initialization step where we take n_0 samples from each system (i, j) and calculate the sample means $\bar{X}_{ij}(n_0)$. Then, the procedure enters a loop. At each stage when the used budget n_{ij} is smaller than the given total budget N, the procedure recommends a system based on all the past samples and then simulates one additional sample from this system. When the total budget is exhausted, the procedure terminates and returns the best alternative.

Procedure 1: ROBUST UCB PROCEDURE

(1) Setup:

Given the total budget N and determine the first-stage sample size n_0.

(2) Initiation:

Generate n_0 samples from each system (i, j) and calculate its sample mean $\bar{X}_{ij}(n_0)$.

Let $\mathrm{n}_{ij,t} = \mathrm{n}_0$ and $t = kmn_0$.

(3) Inner-layer Selection:

For each alternative $s_i \in S$, choose the probability scenario j_t^i with the largest upper confidence bound, which satisfies

$$\mathrm{j}_t^i = \operatorname*{argmax}_j \left(\bar{X}_{ij}(n_{ij,t}) + \sigma_{ij} \sqrt{\frac{2\log n_{i.,t}}{n_{ij,t}}} \right),$$

where $n_{i.,t}$ is the total budget allocated to all systems related to alternative s_i.

(4) Outer-layer Selection:

Choose the alternative i_t which satisfies

$$\mathrm{i}_t = \operatorname*{argmin}_i \left(\bar{X}_{ij_t^i}\left(n_{ij_t^i,t}\right) - \sigma_{ij_t^i} \sqrt{\frac{2n_t^*}{n_{ij_t^i,t}}} \right), \tag{2}$$

where n_t^* is the total budget allocated to the worst-case probability scenarios of all alternatives.

(5)Update:

Generate one sample from system (i_t, j_t) and update its sample mean. Set $\mathrm{n}_{i_t j_t, t+1} = \mathrm{n}_{i_t j_t, t} + 1$ and $t = t + 1$.

(6)Stopping Criterion

When $\sum_{i=1}^{k}\sum_{i=1}^{m} n_{ij,t} \geq N$, the procedure is terminated. Return alternative

$$\underset{i=1,2,\ldots,k}{\arg\min}\left(\max_{j=1,2,\ldots,m} \bar{X}_{ij}(n_{ij,N})\right)$$

as the best alternative.

3.4 Consistency

In this part, we prove the consistency of the R-UCB procedure. Amemiya (1985) points out that when the sample size increases indefinitely, the estimator will converge to the true value, which is called consistency.

Theorem 1 When the total budget N goes to infinity, the R-UCB procedure will select the true best alternative with probability 1, i.e., $PCS \xrightarrow{a.s} 1$ and EOC $\rightarrow 0$ as $N \rightarrow \infty$.

Proof. To prove Theorem 1, it suffices to prove the following statement: As N goes to infinity, the budget allocated to each system (i, j) also goes to infinity. It is because, as long as each system is sampled infinite times, by the Strong Law of Large Numbers (Loève 1977), the sample mean of each system will ultimately converge to its true mean. This implies that, in the limit as $N \rightarrow \infty$, the procedure will select the true best with probability one and accordingly EOC goes to zero.

Notice that the statement above is guaranteed whenever the following two statements are satisfied:

- **Statement 1 (Outer-layer budget allocation)**: When N goes to infinity, the total budget allocated to all m systems of alternative s_i, i.e., $n_{i\cdot,N} = \sum_{j=1}^{m} n_{ij,N}$, goes to infinity .
- **Statement 2 (Inner-layer budget allocation)**: When $n_{i\cdot,N}$ goes to infinity for each alternative s_i, the budget allocated to each system (i, j), i.e., $n_{ij,N}$, goes to infinity.

Because the justifications for Statements 1 and 2 are nearly the same, we only include the part for Statement 2 in this proof. Particularly, we justify Statement 2 by contradiction. Suppose that, as $N \rightarrow \infty$, the total simulation budget $n_{i\cdot,N}$ allocated to the systems of alternative s_i also goes to infinity but there exists some system (i, j) which only receives a finite simulation budget. Then, according to the inner-layer sample allocation policy stated in equation (2), the upper confidence bound of system (i, j) would become infinite. In this situation, system (i, j) will consistently receive samples until its sample size grows to infinity so that the upper confidence bound becomes finite. Clearly, this leads to a contradiction. Therefore, Statement 2 is justified.

□

4 NUMERICAL EXPERIMENTS

In this section, we compare the R-UCB procedure with the existing ROCBA procedure by performing a series of numerical experiments. We consider the slippage configurations (SC) of the means, i.e.,

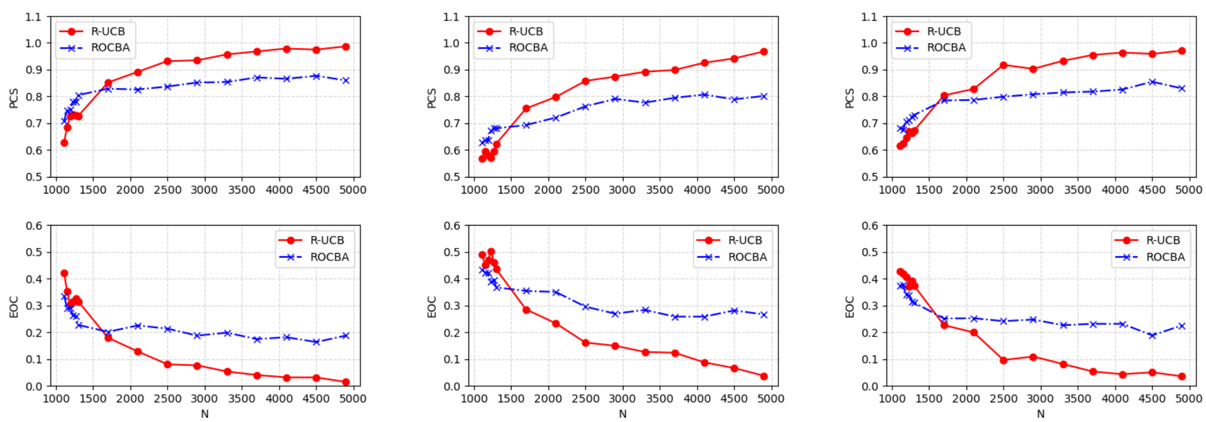
$$[\mu_{ij}]_{k\times m} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 3 & 2 & \ldots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 3 & 2 & \ldots & 2 \end{pmatrix}.$$

which is known as the least favorable configuration in the classic R&S literature. The best alternative is defined as the one with $\min_i \max_j \bar{X}_{ij}$. In other words, alternative 1 is the best alternative. Besides, we consider three configurations of variances:

(1) equal variances (EV) with $\sigma_{ij} = 8$;
(2) increasing variances (IV) with $\sigma_{ij} = 8 + 0.2 * (j-1)$;
(3) decreasing variances (DV) with $\sigma_{ij} = 8 - 1/(1 + 0.2 * (j-1))$.

In this section, we would like to investigate how the performances of both procedures are affected by the total budget N, the number of distributions in the ambiguity set m and the number of alternatives k. For each combination of N, k, m, we conduct 1000 macro-replications to estimate the PCS and EOC of both procedures. The first-stage sample size is set as $n_0 = 10$.

First of all, we test the effect of budget N on the performance of both procedures and present the results in Figure 1. For each configuration of variances, we see that the PCS and EOC achieved by both procedures exhibit the same trend. In particular, for both procedures, the PCS (EOC) first increases (decreases) to a certain level and then remains stable as the total budget N grows. Besides, while the ROCBA procedure performs better than our R-UCB procedure when the budget is relatively small, the R-UCB procedure surpasses the ROCBA procedure when the sample size reaches approximately 1500. As N further grows, the PCS of R-UCB converges to nearly 100% whereas that of the ROCBA reaches a level of about 80%.



(a) EV (b) IV (c) DV

Figure 1: PCS and EOC of the R-UCB and ROCBA procedures with varying budget N.

Figure 2 shows the influence of the size of ambiguity set, m, in which the variances are set under EV. The results for IV and DV are similar and thus we omit them due to the page limit. In this experiment, we set the total budget as $N = (2n_0) \times km$ which grows linearly in m. From Figure 2, it can be found that the PCS increases and the EOC decreases as m increases. Notice that, by the setting of N, the inner-layer problem of RSB is allocated to a larger simulation budget as m increases, which consequentially leads to a more accurate estimation for the inner-layer worst-case mean performance and further a higher PCS (or a lower EOC). Meanwhile, the R-UCB procedure attains a higher PCS and a lower EOC than ROCBA when m is large. This indicates that, the R-UCB procedure can be served as a useful tool to address the RSB problem with a relatively large size of ambiguity set.

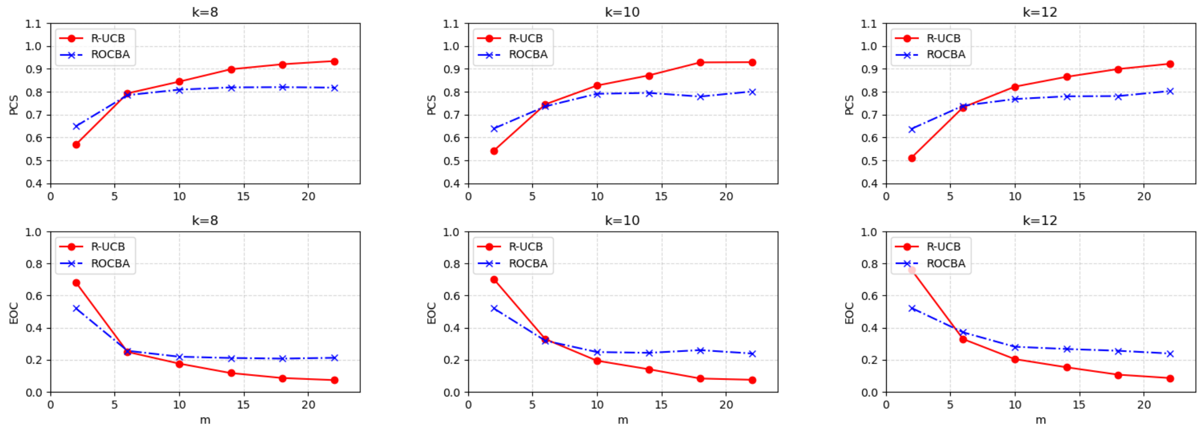


Figure 2: PCS and EOC of the R-UCB and ROCBA procedures with varying size of ambiguity set.

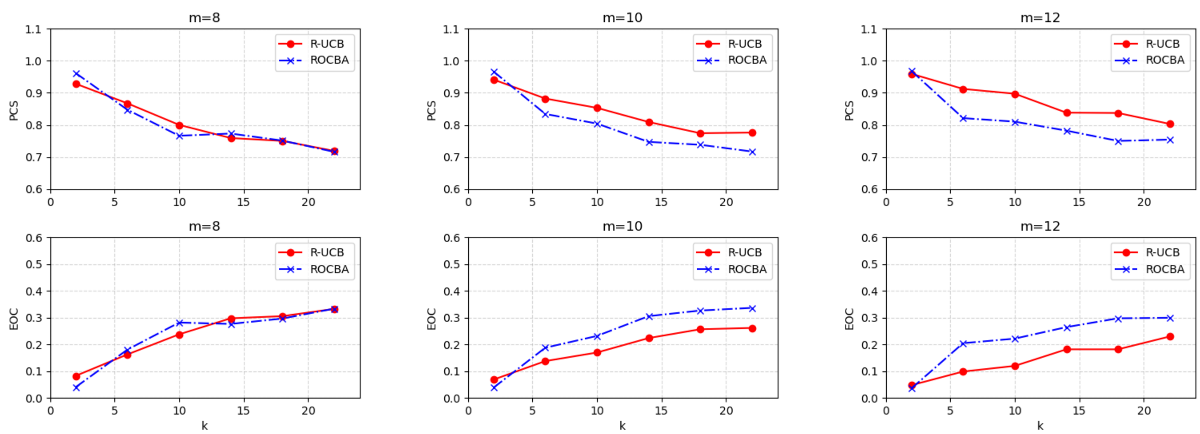


Figure 3: PCS and EOC of the R-UCB and ROCBA procedures with varying numbers of alternatives k.

Figure 3 investigates how the change of k affects the PCS and EOC of both procedures. Following the setting of the simulation budget N in Figure 2, we set $N = (2n_0) \times km$. In contrast with Figure 2, the PCS (EOC) of both procedures reduces (grows) quickly as k increases. It is because, as the number of alternatives is enlarged, it becomes harder for procedures to select the best alternative correctly. In addition, when k is small, the R-UCB has a similar performance with the ROCBA. However, as k grows, R-UCB keeps enjoying a superior performance, i.e., a higher PCS and a lower EOC, than the ROCBA. This indicates that, the R-UCB might utilize a more efficient sampling allocation policy when k is moderate or large.

5 CONCLUSION

This paper studies the R&S problem under input uncertainty, which is called RSB in the literature. Our goal is to propose a new fixed-budget procedure for solving this problem. Unlike traditional RSB procedures, our new procedure is based on the famous UCB procedure from the field of MAB, for its simplicity and theoretical guarantee, and we call it the R-UCB procedure. Due to the two-layer structure of the RSB problem, the R-UCB procedure naturally consists of two UCB procedures, one for the inner-layer problem and the other for the outer-layer problem. As the objectives in the two layers are different from that of the traditional UCB procedure, we carefully modify the exploration rate of the traditional UCB procedure so that they can meet the objective in each layer. We prove that the new procedure is consistent and preliminary numerical experiments support its effectiveness.

ACKNOWLEDGEMENT

This work was partially supported by the National Natural Science Foundation of China [No. 72091211, 72161160340 and 72071146].

REFERENCES

Amemiya, T. 1985. *Advanced Econometrics*. 1st ed. Cambridge, MA: Harvard University Press.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath. 1999. "Coherent Measures of Risk". *Mathematical Finance* 9(3):203–228.

Auer, P., N. Cesa-Bianchi, Y. Freund, and R. E. Schapire. 2002. "The Nonstochastic Multiarmed Bandit Problem". *SIAM Journal on Computing* 32(1):48–77.

Bechhofer, R. E. 1954. "A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances". *Annals of Mathematical Statistics* 25(1):16–39.

Bubeck, S., R. Munos, and G. Stoltz. 2009. "Pure Exploration in Multi-Armed Bandits Problems". In *Proceedings of the 20th International Conference on Algorithmic Learning Theory*. October 3rd-5th, Porto, Portugal, 23-27.

Fan, W., L. J. Hong, and X. Zhang. 2013. "Robust Selection of the Best". In *Proceedings of the 2013 Winter Simulation Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. R. Hill, and M. E. Kuhl, 868–876. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Fan, W., L. J. Hong, and X. Zhang. 2020. "Distributionally Robust Selection of the Best". *Management Science* 66(1):190–208.

Gabillon, V., M. Ghavamzadeh, and A. Lazaric. 2012. "Best Arm Identification: A Unified Approach to Fixed Budget and Fixed Confidence". In *Proceedings of the 25th International Conference on Neural Information Processing Systems*. December 3rd-6th, Lake Tahoe, Nevada, 3212–3220.

Gao, S., H. Xiao, E. Zhou, and W. Chen. 2017. "Robust Ranking and Selection with Optimal Computing Budget Allocation". *Automatica* 81:30–36.

Hong, L. J., W. Fan, and J. Luo. 2021. "Review on Ranking and Selection: A New Perspective". *Frontiers of Engineering Management* 8(3):321–343.

Hunter, S. R., and B. L. Nelson. 2017. "Parallel Ranking and Selection". In *Advances in Modeling and Simulation*, edited by A. Tolk, J. Fowler, G. Shao, and E. Yücesan, 249–275. New York: Springer.

Lesnevski, V., B. L. Nelson, and J. Staum. 2007. "Simulation of Coherent Risk Measures Based on Generalized Scenarios". *Management Science* 53(11):1756–1769.

Liu, G., W. Shi, and K. Zhang. 2019. "An Upper Confidence Bound Approach to Estimating Coherent Risk Measures". In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H. G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, 914–925. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Loève, M. 1977. *Probability Theory*. 4th ed. New York: Springer.

Zhang, X., and L. Ding. 2016. "Sequential Sampling for Bayesian Robust Ranking and Selection". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 758–769. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

YUCHEN WAN is a Ph.D. student in the School of Data Science at Fudan University in Shanghai, China. She received her bechelor's and master's degree from Soochow University and the University of Warwick, respectively. Her research interests include ranking and selection, and systemic financial risk management. Her email address is 22110980026@m.fudan.edu.cn.

WEIWEI FAN is an associate professor in the Advanced Institute of Business and School of Economics and Management at Tongji University in Shanghai, China. She received her PhD in the Department of Industrial Engineering and Logistics Management at Hong Kong University of Science and Technology. Her research interests include simulation optimization, robust optimization and their applications in healthcare management and supply chain management. Her email address is wfan@tongji.edu.cn.

L. JEFF HONG is the Fudan Distinguished Professor and Hongyi Chair Professor with joint appointment at School of Management and School of Data Science at Fudan University in Shanghai, China. His research interests include stochastic simulation, stochastic optimization, risk management and supply chain management. He is currently the simulation area editor of *Operations Research* and an associate editor of *Management Science* and *ACM Transactions on Modeling and Computer Simulation*. His email address is hong_liu@fudan.edu.cn.