Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

GENERAL-PURPOSE RANKING AND SELECTION FOR STOCHASTIC SIMULATION WITH STREAMING INPUT DATA

Jaime Gonzalez-Hodar[1]

[1]School of Industrial and Systems Eng., Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT

We study ranking and selection (R&S) where the simulator's input models are more precisely estimated from the streaming data obtained from the system. The goal is to decide when to stop updating the model and return the estimated optimum with a probability of good selection (PGS) guarantee. We extend the general-purpose R&S procedure by Lee and Nelson by integrating a metamodel that represents the input uncertainty effect on the simulation output performance measure. The algorithm stops when the estimated PGS is no less than $1-\alpha$ accounting for both prediction error in the metamodel and input uncertainty. We then propose an alternative procedure that terminates significantly earlier while still providing the same (approximate) PGS guarantee by allowing the performance measures of inferior solutions to be estimated with lower precision than those of good solutions. Both algorithms can accommodate nonparametric input models and/or performance measures other than the means (e.g., quantiles).

1 INTRODUCTION

We consider an R&S problem where the goal is to identify the best system among k alternatives and the performance measures are estimated from outputs generated by a stochastic simulator. Formally, the objective is to find $i^c \triangleq \text{argmax}_{1\le i\le k}\, \eta_i(\boldsymbol{F}^c)$, where $\boldsymbol{F}^c \triangleq \{F_1^c, \dots, F_L^c\}$ is the collection of input distributions shared across all systems and $\eta_i(\boldsymbol{F})$ denotes the performance measure of system i when the inputs are generated from $\boldsymbol{F}$. We assume $\boldsymbol{F}^c$ is unknown but estimated from data. Consequently, in addition to stochastic simulation noise, the outputs are also subject to uncertainty that comes from the estimated input models, known as *input uncertainty*.

Specifically, we focus on the problem when input models are periodically updated from streaming data collected over multiple periods. We assume that, for each $p \in \{1, \dots, L\}$, a streaming process generates independent and identically distributed observations, which are collected in batches at the end of each period. Then, we update the estimator of $\boldsymbol{F}^c$, denoted as $\widehat{\boldsymbol{F}} \triangleq \{\hat{F}_1, \dots, \hat{F}_L\}$, using all collected data. Our framework lets the user decide whether to stop the data collection and return the current estimated optimum after each period. This decision is based on the PGS exceeding $1-\alpha$, where the PGS is defined as the probability that the selected system $\hat{\imath}$ is within a user-defined δ difference from the best.

The main contribution of this work is three-fold. First, to the best of our knowledge, our framework is the first to tackle R&S under input uncertainty completely nonparametrically. Second, our framework can accommodate general performance measures other than the mean such as quantiles. Lastly, our procedure provides an asymptotic PGS for an R&S problem under input uncertainty.

2 METAMODELING

To design an efficient procedure, we adopt a metamodel to predict $\eta_i(\boldsymbol{F})$ at arbitrary input model $\boldsymbol{F}$ without having to run simulations at $\boldsymbol{F}$. We model $\eta_i(\boldsymbol{F})$ as a function of the moments of $\boldsymbol{F}$. We denote the moment vector by $\boldsymbol{\theta} \triangleq (\theta_1, \dots, \theta_D)^\top$ and impose a linear model in $\boldsymbol{\theta}$, $\psi_i(\boldsymbol{\theta}) \triangleq \boldsymbol{\theta}^\top \boldsymbol{\beta}^i$. This can be generalized by introducing basis functions.

We bootstrap the observations B times to obtain $\widehat{\boldsymbol{F}}^{*(1)}, \dots, \widehat{\boldsymbol{F}}^{*(B)}$ (and so $\widehat{\boldsymbol{\theta}}^{*(1)}, \dots, \widehat{\boldsymbol{\theta}}^{*(B)}$) and run simulations at each. Then, the least squares estimator of $\boldsymbol{\beta}^i$ and the fitted linear model for arbitrary $\boldsymbol{\theta}$ are $\widehat{\boldsymbol{\beta}}^i = (\boldsymbol{\Theta}^\top \boldsymbol{\Theta})^{-1} \boldsymbol{\Theta}^\top \boldsymbol{y}_i$ and $\hat{\psi}_i(\boldsymbol{\theta}) = \boldsymbol{\theta}^\top \widehat{\boldsymbol{\beta}}^i$, where $\boldsymbol{\Theta}$ is the $B \times D$ matrix whose bth row is $\widehat{\boldsymbol{\theta}}^{*(b)}$ and $\boldsymbol{y}_i$ is the vector of simulation outputs. Under some simplifying assumptions, we also characterize the prediction error of $\hat{\psi}_i(\boldsymbol{\theta})$ to be normally distributed with mean zero and variance $\sigma_i^2 (\boldsymbol{\Theta}^\top \boldsymbol{\Theta})^{-1}$.

Continuously refitting the model with all cumulative data can be computationally expensive. We introduce a parameter $\gamma \in (0,1]$ to control the fraction of periods whose bootstrapped design points and simulation outputs are included in the regression. This parameter also helps balancing the bias in the metamodel produced by initial estimations of $\boldsymbol{F}^c$ and the variance coming from the estimation error.

3 BOOTSTRAP PROCEDURES

We extend the general-purpose R&S procedure proposed by Lee and Nelson (2016), by integrating the metamodel with bootstrapping. Hsu (1996) shows that, if we have an estimator $\hat{\eta}_i(F^c)$ of $\eta_i(F^c)$ such that

$$\Pr\left\{\hat{\eta}_i(\boldsymbol{F}^c) - \hat{\eta}_j(\boldsymbol{F}^c) - \left(\eta_i(\boldsymbol{F}^c) - \eta_j(\boldsymbol{F}^c)\right) \leq \delta, \forall i \neq j\right\} \geq 1 - \alpha, \tag{1}$$

then selecting $\hat{\imath} = \operatorname{argmax}_{1 \leq i \leq k} \hat{\eta}_i(\boldsymbol{F}^c)$ as the best ensures PGS $\geq 1 - \alpha$. In our case, since $\boldsymbol{F}^c$ is unknown, we replace $\hat{\eta}_i(\boldsymbol{F}^c)$ with $\hat{\psi}_i(\widehat{\boldsymbol{\theta}})$, where $\widehat{\boldsymbol{\theta}}$ is the vector of moments of $\widehat{\boldsymbol{F}}$. Thus, the probability in (1) must be taken with respect to the sampling distribution of $\widehat{\boldsymbol{F}}$ as well as the prediction error of the metamodel.

Since $\eta_i(\boldsymbol{F}^c)$ does not have any analytical form, we compute the bootstrap estimator of (1) as

$$\frac{1}{B} \sum_{b=1}^{B} \mathbf{1}\left\{\left(\widehat{\boldsymbol{\theta}}^{*(b)} - \widehat{\boldsymbol{\theta}}\right)^\top \left(\widehat{\boldsymbol{\beta}}^i - \widehat{\boldsymbol{\beta}}^j\right) + \left(\widehat{\boldsymbol{\theta}}^{*(b)} - \widehat{\boldsymbol{\theta}}\right)^\top \left(\boldsymbol{\xi}_i^{(b)} - \boldsymbol{\xi}_j^{(b)}\right) \leq \delta, \forall i \neq j\right\}, \tag{2}$$

where $\mathbf{1}\{\cdot\}$ is the indicator function and $\boldsymbol{\xi}_i^{(b)}$ is a random vector which characterizes the prediction error in the metamodel. In the full paper, we propose Algorithm 1 that computes (2) after collecting additional data in every period and decides if more data is needed to guarantee the desired PGS. With the introduction of the parameter γ, the prediction error also decreases over time.

We formulate Algorithm 2 that improves the efficiency over Algorithm 1. Instead of (1), we guarantee the PGS using a less conservative probability bound (Nelson and Banerjee, 2001):

$$\Pr\left\{\hat{\eta}_i(\boldsymbol{F}^c) - \hat{\eta}_{i^c}(\boldsymbol{F}^c) - \left(\eta_i(\boldsymbol{F}^c) - \eta_{i^c}(\boldsymbol{F}^c)\right) \leq \max\{\delta, \eta_{i^c}(\boldsymbol{F}^c) - \eta_i(\boldsymbol{F}^c)\}, \forall i \neq i^c\right\} \geq 1 - \alpha. \tag{3}$$

The probability bound in (3) has two benefits. First, we reduce the number of comparisons from $k(k-1)$ to $k-1$ by focusing on the comparisons between each solution with (the estimated) i^c. Second, the comparisons between i^c and considerably inferior systems are allowed to be less precise.

4 RESULTS

We test both algorithms in two synthetic examples where we can model the performance measure as a function of the first moment of the input distributions. The third is a queueing system example. Empirical results suggest that both algorithms provide a valid PGS. In addition, Algorithm 2 reduces the input data requirement by up to 90% compared to Algorithm 1, resulting in earlier termination.

REFERENCES

Hsu, J. C. 1996. Multiple Comparisons: Theory and Methods. 1st ed. London: Chapman & Hall/CRC.

Lee, S., and B. L. Nelson. 2016. "General-Purpose Ranking and Selection for Computer Simulation". IIE transactions 48(6):555-564.

Nelson, B. L., and S. Banerjee. 2001. "Selecting a Good System: Procedures and Inference". IIE Transactions 33(3):149–166.