Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo and M. Loper, eds.

CONTEXTUAL RANKING AND SELECTION WITH GAUSSIAN PROCESSES

Sait Cakmak

School of Industrial and Systems Engineering
Georgia Institute of Technology
755 Ferst Drive NW
Atlanta, GA 30332, USA

ABSTRACT

In many real world problems, we are faced with the problem of selecting the best among a finite number of alternatives, where the best alternative is determined based on context specific information. In this work, we study the contextual Ranking and Selection problem under a finite arm - finite context setting, where we aim to find the best alternative for each context. We use a separate Gaussian process to model the reward for each arm, derive the large deviations rate function for both the expected and worst-case contextual probability of correct selection, and propose an iterative algorithm for maximizing the rate function. Numerical experiments show that our algorithm is highly competitive in terms of sampling efficiency, while having significantly smaller computational overhead.

1 INTRODUCTION

Ranking & Selection (R&S) studies the problem of identifying the best among a finite number of alternatives (arms), where the true performance of each alternative is only observed through noisy evaluations. The settings of R&S can be typically categorized into fixed confidence and fixed budget. In the fixed-confidence setting, the goal is to achieve a target probability of correct selection (PCS) of the best alternative using as few evaluations as possible, while in the fixed-budget setting one aims to achieve a PCS as high as possible with the given sampling budget.

In certain applications, the best alternative may not be the same across the board. The performance of an alternative may depend environmental variables that are beyond the control of the decision maker, called *contexts*. The benefit of making context-dependent decisions is easily seen by a simple application of Jensen's inequality: $E_c[max_k\, f(k,c)] \geq max_k E_c[f(k,c)]$, where $f(k,c)$ represents the reward of selecting alternative k for the context c, and $E_c[\cdot]$ denotes the expectation with respect to (w.r.t.) c. Examples of context-dependent decision making include personalized medicine, where the best drug and dose may depend on the patient's age, gender, and medical history; and a reconfigurable manufacturing system, where based on a set of forecasted market conditions, a set of alternative system configurations can be simulated to determine the most profitable configuration to use when the market conditions are realized.

In this work, we study the contextual R&S problem, where the goal is to identify the best alternative for each context under a fixed budget. At each iteration, the decision maker selects an arm-context pair to evaluate and observes a noisy evaluation of the reward function. Since the best arm cannot be identified with certainty using the finite sampling budget, we aim to design an adaptive sampling policy that maximizes an "aggregated" (expected or worst-case) PCS over contexts, referred to as contextual PCS.

2 RELATED WORKS

The contextual R\&S problem has seen an increasing interest in past few years. Li et al. (2020) focus on worst-case PCS, use independent normal random variables to model rewards, and propose a one-step lookahead policy with an efficient value function approximation. Shen et al. (2021) assume that the reward for each arm is a linear function of the context and propose a two-stage algorithm based on the indifference zone formulation for optimizing the expected PCS. The most closely related work to ours is Gao et al. (2019), which extends the analysis in Glynn and Juneja (2004) to derive the large deviations rate function for the contextual PCS under a simple frequentist model and proposes an adaptive sampling policy that allocates samples one-by-one to maximize the rate function.

3 METHODOLOGY & RESULTS

In this work, we use a separate Gaussian process (GP) to model the reward function for each arm. By leveraging the hidden correlation structure within the reward function, GPs offer significant improvements in posterior inference over independent normal random variables, which are commonly used in the R&S literature, including in Gao et al. (2019). Using the GP posterior mean as the predictor of the rewards, we derive the large deviations rate function of the contextual PCS, and propose the GP-C-OCBA sampling policy, which uses the GP posterior mean and variance to determine the next arm-context to sample from.

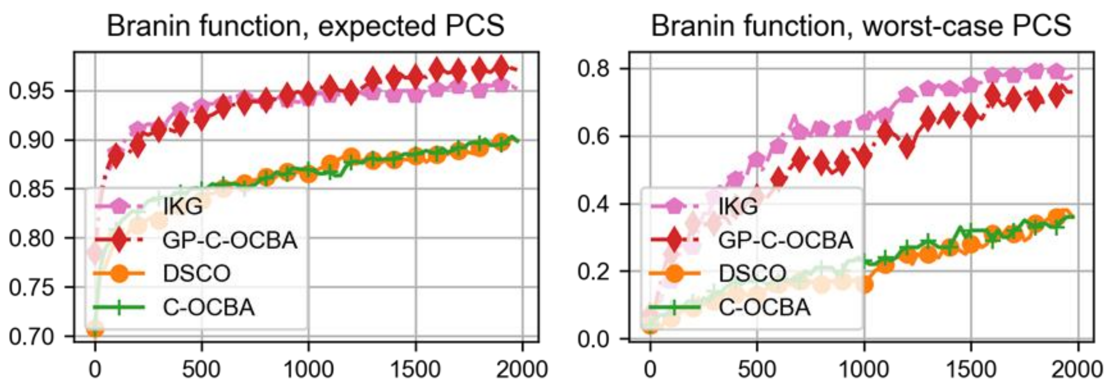

Figure 1: Expected and worst-case PCS achieved by algorithms on the Branin function.

Numerical experiments, partly presented in Figure 1, show that GP-C-OCBA is highly sample efficient, and requires significantly lower computational overhead than the integrated knowledge gradient (IKG, Pearce & Branke, 2018), the only other competitive algorithm in our experiments. The minimum and maximum average runtime for running 1000 iterations of the experiments was 1218 and 43224 seconds for IKG, which is significantly slower than 249 and 544 seconds it took to run GP-C-OCBA.

REFERENCES

Gao, S., J. Du, and C. Chen. 2019. "Selecting the Optimal System Design under Covariates". In 15th International Conference on Automation Science and Engineering. August 22nd-26th, Vancouver, BC, Canada, 547-552.

Glynn, P., and S. Juneja. 2004. "A Large Deviations Perspective on Ordinal Optimization". In Proceedings of the 2004 Winter Simulation Conference, edited by R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 577–585. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Li, H., H. Lam, Z. Liang, and Y. Peng. 2020. "Context-Dependent Ranking and Selection under a Bayesian Framework". In Proceedings of the 2020 Winter Simulation Conference, edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, 2060–2070. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Pearce, M., and J. Branke. 2018. "Continuous Multi-Task Bayesian Optimisation with Correlation". European Journal of Operational Research 270(3): 1074-1085.

Shen, H., L. J. Hong, and X. Zhang. 2021. "Ranking and Selection with Covariates for Personalized Decision Making". INFORMS Journal on Computing. Articles in advance, published online.