Cost-Aware Active Learning for Feasible Region Identification

Ioana Nikova
ioana.nikova@siemens.com
Siemens Industry Software
Leuven, Belgium
Ghent University
Ghent, Belgium

Tom Dhaene
Ivo Couckuyt
tom.dhaene@ugent.be
ivo.couckuyt@ugent.be
Ghent University - imec
Ghent, Belgium

ABSTRACT

Design space exploration for engineering design involves identifying feasible designs that satisfy design specifications, often represented by feasibility constraints. To determine whether a design is feasible, an expensive simulation is required. Therefore, it is crucial to find and model the feasible region with as few simulations as possible.

Model-based Active learning (AL) is a data-efficient, iterative sampling framework that can be used for design space exploration to identify feasible regions with the least amount of budget spent. A common choice for the budget is the number of (sampling) iterations. This is a good choice when every simulation has an equal cost. However, simulation cost can vary depending on the design parameters and is often unknown. Thus, some regions in the design space are cheaper to evaluate than others.

In this work, we investigate if incorporating the unknown cost in the AL strategy leads to better sampling and, eventually, faster identification of the feasible region.

CCS CONCEPTS

• **Computing methodologies** → **Active learning settings**; • **Mathematics of computing** → **Bayesian computation**.

KEYWORDS

design space exploration, active learning, Gaussian Process, feasible region

ACM Reference Format:

Ioana Nikova, Tom Dhaene, and Ivo Couckuyt. 2023. Cost-Aware Active Learning for Feasible Region Identification. In *Genetic and Evolutionary Computation Conference Companion (GECCO '23 Companion), July 15–19, 2023, Lisbon, Portugal.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3583133.3596323

1 INTRODUCTION

Consider a typical engineering problem where the engineer has to search over a design space $\mathcal{X} \in \mathbb{R}^d$ to find an optimal solution. Early in the design process, the engineer first explores the design space (i.e., instead of optimization) to find distinct and feasible design solutions. The design requirements are specified by a set of L constraints: $G(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_L(\mathbf{x}))^\top \le (t_1, \dots, t_L)^\top$, where $g_l : \mathbb{R}^d \to \mathbb{R}$ is the l-th constraint function with threshold t_l.

As running simulations to calculate $G(\mathbf{x})$ is computationally expensive, there is a need for data-efficient sampling techniques for feasible region identification (FRI). Active learning (AL) is a sequential approach that uses a surrogate model, a cheap approximation of a simulation, to intelligently extend the data set $\mathcal{D}$. The sampling strategy of Bayesian AL is defined by an acquisition function $\alpha(\mathbf{x}, \hat{G}, \mathbf{t})$ which is optimized to find the next sample. In the case of FRI, there are model-dependent acquisition functions that sample along the boundary, like *Entropy Feasible (EF)* [4] and *Probability of being at the Boundary and Entropy (PBE)* [6], and others that sample inside the feasible regions, like *Probability of Feasibility and Variance (PoFV)* [3].

A common choice for the surrogate model used by Bayesian AL is a Gaussian Process (GP) [4] as it provides uncertainty information for the prediction. In this work, we model each constraint independently with a standard GP.

The AL strategy stops when a certain budget has been depleted. Often the budget is defined by the number of iterations. This choice is data-efficient and cost-effective when the simulations have an equal cost. However, there exist engineering problems where the simulation cost, measured in time, money, or energy, can vary in the design space.

Research to date has focused on cost-aware Bayesian Optimization methods for hyperparameter optimization. These cost-aware methods either incorporate the known cost in the acquisition function [9] or model the unknown cost [5]. To our knowledge, cost-aware extensions have not been investigated for engineering problems or acquisition functions for FRI.

Our contribution is to extend the PoFV acquisition function with cost-aware strategies by adding a fixed cost component [8] or using cost-cooling [5]. In both cases, the cost is unknown and modeled by a GP. We test this extended approach on a highly-constrained engineering problem with an artificial cost function.

2 COST-AWARE ACTIVE LEARNING ON A REAL-WORLD PROBLEM

We focus on extending the PoFV acquisition function because the cost can also influence the sampling in the feasible region. For problems with multiple constraints, PoFV can be defined as:

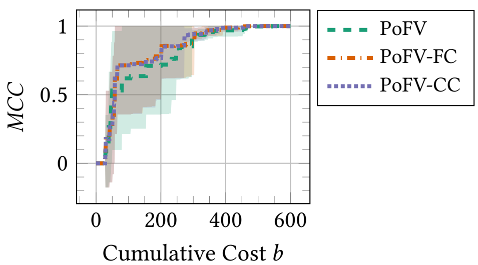
$$\alpha(\mathbf{x}, \hat{G}, \mathbf{t}) = \prod_{l=1}^{L} p\big(p(\hat{g}_l|\mathbf{x}, \mathcal{D}_l) \le t_l\big)\ \sigma_l^2(\mathbf{x}), \tag{1}$$

where we multiply the probability of feasibility by the predictive variance for each constraint.

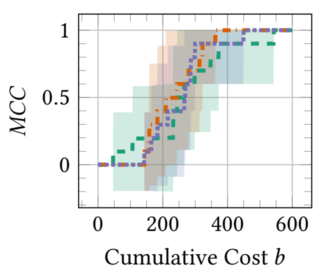
ACM ISBN 979-8-4007-0120-7/23/07.
https://doi.org/10.1145/3583133.3596323

Algorithm 1 CAAL: Cost-Aware Active Learning

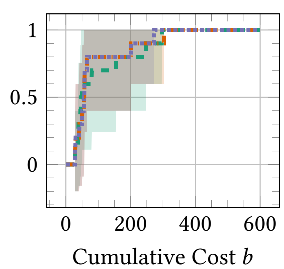
1: **procedure** CAAL(design space $\mathcal{X}$, initial data set size N, constraint models $\hat{G} = \{\hat{g}_1, \ldots, \hat{g}_L\}$, cost model $\hat{c}$, budget β)
2: $\quad \mathcal{D} = G\big(Halton(\mathcal{X}, N)\big)$ ▷ Initial space-filling data set [2]
3: $\quad$ Update $\{\hat{g}_1, \ldots, \hat{g}_L\}$, $\hat{c}$
4: $\quad b \leftarrow 0, \kappa \leftarrow 1$ ▷ Cumulative cost b, cost-weight κ
5: $\quad$ **while** $b \le \beta$ **do**
6: $\quad\quad \mathbf{x}^* \leftarrow argmax_{\mathbf{x}} \frac{\alpha(\mathbf{x}, \hat{G}, t)}{\hat{c}(\mathbf{x})^{\kappa}}$ ▷ α from Eq. 1
7: $\quad\quad$ Extend $\mathcal{D}$ with $G(\mathbf{x}^*)$
8: $\quad\quad$ Update b, $\{\hat{g}_1, \ldots, \hat{g}_L\}$, $\hat{c}$
9: $\quad\quad \kappa \leftarrow \frac{\beta - b}{\beta}$ ▷ Only when cost-cooling
10: $\quad$ **end while**
11: **end procedure**



(a) $\hat{G}(\mathbf{x})$



(b) $\hat{g}_1(\mathbf{x})$



(c) $\hat{g}_7(\mathbf{x})$

Figure 1: Speed Reducer problem: MCC of (a) all constraint models combined; (b) less restrictive constraint model; (c) most restrictive constraint model. The mean and standard deviation of 10 runs are plotted for $\beta = 600$.

We discuss two approaches to adding cost awareness. Firstly, we consider the fixed cost approach by Snoek et al. [8]. The acquisition value is divided by the predicted cost $\hat{c}(\mathbf{x})$. We refer to this method as PoFV-FC. Secondly, we use the cost-cooling approach by Lee et al. [5]. The cost-cooling method further denoted as PoFV-CC uses a cost-weight κ. Hence, the influence of $\hat{c}(\mathbf{x})$ decreases as the AL method progresses. Algorithm 1 shows the cost-aware AL method. Note that the fixed cost approach can be written with the cost-weight, where $\kappa = 1$ at every iteration, and line 9 is skipped.

The cost-aware approaches are tested on an adapted version of Golinski's Speed Reducer problem [7] with seven design variables and seven constraints (see supplementary material A). Experiments show that g_7 is the most restrictive constraint, while g_3, g_4, and g_5 are not restricting the design, so always feasible. The feasibility ratio on a data set of 1000 samples is 56.54%. We chose an exponential cost function as described in the supplementary material.

Figure 1 compares the cost-aware acquisition functions. The models are tested on 1000 samples using the *Matthews Correlation Coefficient (MCC)* [1], a balanced correlation coefficient. Figure 1a combines the performance of all constraint models. All methods give similar results. However, the individual performance depicted in Figures 1b and 1c shows that less restrictive models have been learned better with the cost-aware methods. No improvement is visible for the most restrictive constraint model, which defines the feasible region.

Looking at the final data set, PoFV-CC and PoFV-FC sampled on average 2–3 additional samples, compared to PoFV using the same budget. Of those samples, there is an increase in the number of feasible designs, which the engineer can further use. 77.2% of the PoFV data set is feasible, while 81.9-83.3% is feasible for, respectively, PoFV-CC and PoFV-FC.

3 CONCLUSION

Although the cost-aware methods do not lead to significant model performance improvement, the final data set contains more designs. Furthermore, there is an increase in the number of feasible designs that the engineer can use in the next steps of the design process.

Future research should focus on testing the approaches on more real-world problems with artificial and actual cost functions. On top of that, the cost-aware extensions can also be combined with other acquisition functions.

ACKNOWLEDGMENTS

Ioana Nikova gratefully acknowledges the financial support from the Flemish Agency on Innovation and Entrepreneurship (VLAIO) for the Baekeland Research Project HBC.2021.0841. This work has also been supported by the Flemish Government under the Flanders AI Research Program.

REFERENCES

[1] Sabri Boughorbel, Fethi Jarray, and Mohammed El-Anbari. 2017. Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. *PloS one* 12, 6 (2017), e0177678.
[2] John H. Halton and G. B. Smith. 1964. Algorithm 247: Radical-inverse quasi-random point sequence. *Commun. ACM* 7 (1964), 701–702.
[3] Arun Kaintura, Kyle Foss, Ivo Couckuyt, Tom Dhaene, Odysseas Zografos, Adrien Vaysset, and Bart Sorée. 2018. Machine learning for fast characterization of magnetic logic devices. In *2018 IEEE electrical design of advanced packaging and systems symposium (EDAPS)*. IEEE, 1–3.
[4] Nicolas Knudde, Ivo Couckuyt, Kohei Shintani, and Tom Dhaene. 2019. Active learning for feasible region discovery. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*. IEEE, 567–572.
[5] Eric Hans Lee, Valerio Perrone, Cedric Archambeau, and Matthias Seeger. 2020. Cost-aware Bayesian optimization. *arXiv preprint arXiv:2003.10870* (2020).
[6] Alma Rahat and Michael Wood. 2020. On Bayesian search for the feasible space under computationally expensive constraints. In *Machine Learning, Optimization, and Data Science: 6th International Conference, LOD 2020, Siena, Italy, July 19–23, 2020, Revised Selected Papers, Part II 6*. Springer International Publishing, 529–540.
[7] Tapabrata Ray. 2003. Golinski's speed reducer problem revisited. *AIAA journal* 41, 3 (2003), 556–558.
[8] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. 2012. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems* 25 (2012).
[9] Kevin Swersky, Jasper Snoek, and Ryan P Adams. 2013. Multi-task bayesian optimization. *Advances in neural information processing systems* 26 (2013).

A GOLINSKI'S SPEED REDUCER PROBLEM

For this work, we adapt the original Speed Reducer problem [7] to have seven constraints from the original eleven. The constraints and variables are as follows:

$$g_1(\mathbf{x}) = 27 - x_1 x_2^2 x_3 \leq 0,$$
$$g_2(\mathbf{x}) = 1.93 - \frac{x_2 x_6^4 x_3}{x_4^3} \leq 0,$$
$$g_3(\mathbf{x}) = 1.93 - \frac{x_2 x_7^4 x_3}{x_5^3} \leq 0,$$
$$g_4(\mathbf{x}) = x_2 x_3 - 40 \leq 0,$$
$$g_5(\mathbf{x}) = \frac{x_1}{x_2} - 12 \leq 0,$$
$$g_6(\mathbf{x}) = 1.5x_6 - x_4 + 1.9 \leq 0,$$
$$g_7(\mathbf{x}) = 1.1x_7 - x_5 + 1.9 \leq 0,$$

where $2.6 \leq x_1 \leq 3.6$, $0.7 \leq x_2 \leq 0.8$, $17 \leq x_3 \leq 28$, $7.3 \leq x_4, x_5 \leq 8.3$, $2.9 \leq x_6 \leq 3.9$, and $5 \leq x_7 \leq 5.5$.

The artificial cost function was chosen based on the variables that define the most restrictive constraint g_7. By choosing the variables x_5 and x_7 we want to influence the learning of the most restrictive constraint with the cost. Hence, the exponential cost function used in the experiments is:

$$c(\mathbf{x}) = \frac{e^{x_7}}{2x_5}.$$