Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

A MACHINE LEARNING-AUGMENTED OPTIMIZATION FOR A MAXIMAL COVERING LOCATION-ALLOCATION PROBLEM

Kuangying Li[1], Hiruni Niwunhella[2], Leila Hajibabai[2], and Ali Hajbabaie[3]

[1]Operations Research Program, North Carolina State University, Raleigh, NC, USA
[2]Dept. of Industrial and Syatems Eng., Carolina State University, Raleigh, NC, USA
[3]Dept. of Civil, construction, and Environmental Eng., Carolina State University, Raleigh, NC, USA

ABSTRACT

This study tackles the maximal covering location-allocation problem (MCLP) by introducing a novel machine learning-enhanced column generation (ML-CG) method comprising offline training and online prediction. The proposed model is applied to a vaccine distribution case study, formulated as a mixed-integer linear model to minimize the total cost of vaccine shipments and maximize allocations, incorporating equity constraints for age, race, and gender groups. Empirical case studies in Pennsylvania, based on real-world data from the Centers for Disease Control and Prevention (CDC) and health department websites, confirm the model's effectiveness.

1 INTRODUCTION

Maximal covering location-allocation problems aim to optimize facility placement in a network to maximize demand coverage, a task that becomes particularly challenging in large-scale settings. In previous work, we developed a hybrid approach that integrates Lagrangian relaxation (LR) with a modified Voronoi Diagram to find near-optimal solutions for COVID-19 vaccine distribution at the state level. This hybrid approach demonstrated a 5.92% reduction in total transportation costs and a 28.15% increase in demand coverage compared to the CG method. Despite the promising results, both methods encountered significant computational challenges, leading to extended CPU times. The primary source of this computational burden is the combinatorial complexity of the mixed-integer decision variables in the model and the large-scale nature of its implementation. One potential strategy for improving CPU efficiency is to enhance the integrated LR-Voronoi algorithm. However, the efficiency of the Voronoi diagram can be influenced by various factors, including the specific application and the quality of the input data. Unfortunately, the methods that increase the efficiency of a conventional Voronoi Diagram do not apply to our modified approach due to the different reward functions and strategies employed to reshape sub-regions. Another strategy for enhancing CPU efficiency is to improve the CG technique that served as a benchmark in our prior work. Traditionally, CG is used to solve large-scale mathematical optimization problems by working with a strategically selected subset of variables, forming the restricted master problem (RMP), which is then optimized. Variables that can improve the objective function are added iteratively until no further improvement is achievable. Enhancing the computational efficiency of CG can further be accomplished by incorporating machine learning (ML) techniques to predict key variables. While existing approaches for solving combinatorial optimization problems through the CG method have effectively reduced runtime and ensured better solution quality, their application to our work presents two significant challenges: (i) the combinatorial complexity of the mixed-integer decision variables, with an overwhelming 1.6 billion integral and binary decision variables, given the extensive case study network; and (ii) the problem structure, which involves location-allocation with facility location decisions and allocation quantity decisions. To overcome these challenges, this study proposes an ML-based CG method. Preliminary results indicate that this method effectively solves the problem and outperforms previous solutions by integrating predictive analytics.

2 METHODOLOGY

We have formulated an MCLP as a mixed-integer linear programming (MILP) model. The objective is twofold: to minimize the distribution costs and to maximize vaccine coverage. The model includes allocation decision variables (from distribution centers to providers and from providers to population blocks). Equity constraints ensure that all demographic groups receive an adequate share of vaccines. This study proposes an integrated approach combining LR, CG, and ML to solve the MCLP efficiently. The ML component consists of two phases: (i) Offline Training Phase: A linear support vector machine (SVM) is trained using optimal solutions from previously solved instances of the problem. The training data includes various features representing the problem's constraints and objectives. And, (ii) Online Prediction Phase: The trained SVM model predicts optimal solutions for new instances of the problem in the CG process. This prediction reduces the computational burden by narrowing down the search space for potential solutions. While training the SVM could be computationally intensive depending on the problem size, the proposed approach leverages historical data to train the model offline, thus minimizing the online computational burden. By reducing the search space for potential solutions, the model offers significant computational savings. The ML-CG employs a modified Voronoi diagram technique to handle the spatial distribution of demand points, which is embedded in the LR framework. We have applied the prediction solution to enhance the column-selection strategy in CG.

3 EMPIRICAL ANALYSIS

The empirical analysis is conducted using real-world data from the Centers for Disease Control and Prevention (CDC) and demographic data from Pennsylvania. The analysis first includes data collection and preprocessing: The dataset comprises vaccine allocation records, demographic information, and geographic locations of providers and population blocks. The ML-CG method is applied to optimize vaccine distribution in Pennsylvania. Key metrics, such as runtime, optimality gap, and demand coverage, are used to evaluate the method's performance. The numerical experiments demonstrate that the ML-CG method significantly outperforms traditional CG techniques. The method improves equity in vaccine distribution among different demographic groups. Key results include (i) Runtime Efficiency: The ML-CG method reduces computational time by approximately 84% compared to traditional CG methods; and (ii) Solution Quality: The ML-CG method achieves an optimality gap of just 6.0%, ensuring high-quality solutions.

4 CONCLUSIONS

This study presents an ML-CG method to optimize vaccine allocation. The framework offers an efficient alternative to traditional optimization methods in that it predicts key variables and reduces search space while maintaining a competitive optimality gap. Future research could explore the integration of more advanced ML models, such as convolutional neural networks, to improve the prediction of both continuous and binary variables. Additionally, incorporating real-time data and dynamic models could further enhance the applicability of the methodology. In this study, the data utilized for training and testing the ML model are collected from the CDC and Pennsylvania State Health Department. Our ML model operates deterministically, making predictions based on the same input data without introducing random elements. Once trained, the model follows deterministic rules and generates consistent results for identical inputs. A future direction will be to assess the impact of randomness in the ML-CG procedure considering demand variability over time and geographical locations. During the training phase, the data could be randomly sampled to allow the model to generalize effectively to new instances. For instance, in regions with insufficient data, we plan to generate simulated data and explore performance.