ARTICLE IN PRESS
ARTINT:2847

Artificial Intelligence ••• (••••) •••–•••

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Auction optimization using regression trees and linear models as integer programs

Sicco Verwer[a], Yingqian Zhang[b,*], Qing Chuan Ye[b]

[a] Intelligent Systems Department, Delft University of Technology, The Netherlands
[b] Department of Econometrics, Erasmus University Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received in revised form 14 May 2015
Accepted 21 May 2015
Available online xxxx

Keywords:
Auction design
Machine learning
Optimization
Integer linear programming
Regression

ABSTRACT

In a sequential auction with multiple bidding agents, the problem of determining the ordering of the items to sell in order to maximize the expected revenue is highly challenging. The challenge is largely due to the fact that the autonomy and private information of the agents heavily influence the outcome of the auction.

The main contribution of this paper is two-fold. First, we demonstrate how to apply machine learning techniques to solve the optimal ordering problem in sequential auctions. We learn regression models from historical auctions, which are subsequently used to predict the expected value of orderings for new auctions. Given the learned models, we propose two types of optimization methods: a black-box best-first search approach, and a novel white-box approach that maps learned regression models to integer linear programs (ILP), which can then be solved by any ILP-solver. Although the studied auction design problem is hard, our proposed optimization methods obtain good orderings with high revenues.

Our second main contribution is the insight that the internal structure of regression models can be efficiently evaluated inside an ILP solver for optimization purposes. To this end, we provide efficient encodings of regression trees and linear regression models as ILP constraints. This new way of using learned models for optimization is promising. As the experimental results show, it significantly outperforms the black-box best-first search in nearly all settings.

1. Introduction

One of the main challenges of mathematical optimization is to construct a mathematical model describing the properties of a system. When the structure of a system cannot be fully determined from the knowledge at hand, machine learning and data mining techniques have been used in optimization instead of this knowledge. They have, for example, been used in order to obtain decision values [1], fitness functions [2], or model parameters [3]. Models that have been learned from data are frequently used in a black-box manner, e.g., using only the predictions of learned models but not their internal structure. It is also possible to use these models in a white-box manner, for instance in order to determine search space cuts and parameter bounds. Neural networks have in this way been used to model unknown relations in constraint programming [4]. In this paper, we develop such a white-box optimization method for regression models in integer linear programming, that is, we map these entire models to sets of variables and constraints and solve them using an off the shelf solver.

* Corresponding author.
E-mail addresses: S.E.Verwer@tudelft.nl (S. Verwer), yqzhang@ese.eur.nl (Y. Zhang), ye@ese.eur.nl (Q.C. Ye).

http://dx.doi.org/10.1016/j.artint.2015.05.004

This white-box method together with a proposed black-box method provides a solution to an optimization problem of key interest to the artificial intelligence and operations research communities: auction design. We briefly introduce this problem domain before going into the details of our methods.

1.1. Sequential auction design

Auctions are becoming increasingly popular for allocating resources or items in business-to-business and business-to-customer markets. Often sequential auctions [5] are adopted in practice, where items are sold consecutively to bidders. Sequential auctions are in particular desirable when the number of items for sale is large (e.g., flower auctions [6]), or when the buyers enter and leave the auction dynamically (e.g., online auctions [7]). In a sequential auction, an auctioneer may tune several auction parameters to influence the outcome of an auction, such as reserve prices for items and in which order to sell them. In other words, (s)he can design auctions for the purpose of achieving some predefined goal. In this paper, we solve one specific auction design problem, namely, deciding the optimal ordering of items to sell in a sequential auction in order to maximize the expected revenue (OOSA in short). We assume bidders in such auctions are budget constrained. This is a highly relevant problem in today's auctions since bidders almost always have limited budget, as seen for instance in industrial procurement [8]. Previous research has shown that with the presence of budget constraints, the revenue collected by the auctioneer is heavily dependent on the ordering of items to sell [9–11]. This holds already for a toy problem with 2 items. Let us use a simple example to illustrate the importance of ordering in such cases.

Example 1. Two agents A_1 and A_2 take part in a sequential auction of items. For sale are items r_1 and r_2. Suppose the items are sold by means of first-price, English auction.[1] Assume the reserve prices, which are the lowest prices at which the auctioneer is willing to sell the times, for both items are 1. The amount that agent A_1 and agent A_2 are willing to pay for two items are: $v_1(r_1) = 10$, $v_1(r_2) = 15$, $v_2(r_1) = 12$, $v_2(r_2) = 10$. Furthermore, the budgets of A_1 and A_2 are 15 and 25 respectively.

We assume a simple bidding strategy in this example. The agents bid myopically on each item, that is, their highest bid on one item is the lower value between the amount that they are willing to pay and their remaining budget. The auctioneer's goal is to maximize the total sold price of the items. Consider one situation where the auctioneer sells first r_2 and then r_1. A_1 will get r_2 when she just over-bids A_2 with 11, and then when r_1 is auctioned, A_1 bids maximally 4 due to her budget limit, and A_2 will win the item with the price of 5. The total revenue is 16. However, if the selling sequence is (r_1, r_2), A_2 will win r_1 with the bid 11, and then A_2 will win r_2 with price 11. The collected revenue is 22 in this case. □

Most of the current approaches to the ordering problem in sequential auctions assume a very restricted market environment. They either study the problem of ordering two items, see [11,12], or a market with homogeneous bidders [13]. To the best of our knowledge, we are the first to consider how to order items for realistic auction settings with many heterogeneous bidders competing for many different items. This problem is highly complex—a good design on ordering needs to take care of many uncertainties in the system. For instance, in order to evaluate the revenue given an ordering, the optimization algorithm needs to know the bidders' budgets and preferences on items, which are usually private and unshared. Furthermore, the large variety of possible bidding strategies that bidders may use in auctions are unknown. This auction design problem is a typical example where the mathematical optimization model cannot be fully determined, and hence, machine learning and data mining techniques can come into play. This is exactly what our approach builds upon.

1.2. Learning models for white-box and black-box optimization

Nowadays more and more auctions utilize information technology, which makes it possible to automatically store detailed information about previous auctions along with their selling sequences and the selling price per auctioned item. Our approach to solving the problem of optimal ordering for sequential auctions starts with the historical auction data. We define and compute several relevant features and then use them to learn regression trees and linear regression models for the expected revenue. Given the models, we propose two approaches to find the optimal ordering for a new set of items: (1) a best-first search that uses the models as a black-box to evaluate different orderings of the items; and (2) a novel white-box optimization method that translates the models and the set of items into a mixed-integer program (MIP) and runs this in an ILP-solver (CPLEX). Fig. 1 displays the general framework of our approaches using these two optimization methods.

Just like the traditional black-box optimization approach (see, e.g. [14,15]), our best-first search is ignorant of the internal structure of the models and only calls it to perform function evaluations, i.e., predicting the revenue of an ordering of the items. Optimization is possible by means of a search procedure that uses heuristics to produce new orderings depending on previously evaluated ones. Our best-first search makes use of dynamic programming cuts inspired by sequential decision making in order to reduce the search space.

[1] The English auction that we consider is the one where the starting price is the reserve price, and bidders bid openly against each other. Each subsequent bid should be higher than the previous bid, and the item is sold to the highest bidder at a price equal to her bid.

Please cite this article in press as: S. Verwer et al., Auction optimization using regression trees and linear models as integer programs, Artificial Intelligence (2015), http://dx.doi.org/10.1016/j.artint.2015.05.004