ARTICLE IN PRESS

AUTCON-02037; No of Pages 10

Automation in Construction xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Matching experts' decisions in concrete delivery dispatching centers by ensemble learning algorithms: Tactical level

Mojtaba Maghrebi [a,*], Travis Waller [b], Claude Sammut [c]

[a] School of Civil and Environmental Engineering, University of New South Wales (UNSW), Sydney, Australia and Department Civil Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Iran

[b] School of Civil and Environmental Engineering, University of New South Wales (UNSW), Sydney, Australia

[c] School of Computer Science Engineering, University of New South Wales (UNSW), Sydney, Australia

ARTICLE INFO

Article history:
Received 25 September 2014
Received in revised form 24 February 2016
Accepted 13 March 2016
Available online xxxx

Keywords:
Ready Mixed Concrete (RMC)
Machine learning
Dispatching

ABSTRACT

Ready Mixed Concrete (RMC) suffers from a lack of practical solutions for automatic resource allocation. Under these circumstances, RMC dispatching systems are mostly handled by experts. This paper attempts to introduce a machine learning based method to automatically match experts' decisions in RMC. For this purpose, seven machine learning techniques with their boosted algorithms were selected. A set of attributes was extracted from the collected field data. Eleven metrics were used to assess the performance of the selected techniques using different approaches. Due to concerns about randomness, significant testing was performed to assist in finding the best algorithm for this purpose. Results show that Random-Forest with 85% accuracy outperforms the other selected techniques. One of the most interesting achieved results is related to the computing time. The results show that all the selected algorithms can solve large-scale depot allocations with a very short computing time. This is possibly because a model built by a machine learning algorithm only needs to be tested with new instances, which does not need an extensive computation effort. This provides us with a chance to move toward automation in Ready Mixed Concrete Dispatching Problems (RMCDPs), especially for those RMCs with dynamic environments where resource allocation might need to be quickly recalculated during the RMC process due to changes in the system.

1. Introduction

As was extensively discussed in [1], in an RMC batch plant, based on orders, the specifications of concrete mix are designed and raw materials are mixed together. Then fresh concrete is loaded into a truck. The loaded truck hauls the concrete and pours it at the destination and then returns to the batch plant. In practice, the mixing part is performed automatically; however, the rest of the process is handled by human experts. In detail, dispatchers decide to send a truck from a batch plant at a specific time to a project. This job becomes more complicated when a dispatcher needs to make calculated decisions for supplying concrete for a certain project that is located between two or more batch plants. The dispatcher needs to consider many parameters that can be categorized into three types of information: (i) specification of each order, (ii) travel of truck(s), and (iii) batch plant limitations. Moreover, a Ready Mixed Concrete Dispatching Problem (RMCDP) can be modeled as a network where customers and depots are its nodes and deliveries are the arches between depots and customers. The amount of concrete ordered by different customers is distinct. The number of required deliveries is calculated for each project, based on the ordered amounts.

In the last twenty years, researchers have investigated a variety of approaches to improve the efficiency of RMCDP. However, despite substantial developments in this area, RMCDP still suffers from a lack of practical solutions [2–4] and this process continues to be mostly handled by experts [3,5]. The first drawback of such a human intensive system is its dependence on the human resources, regardless of the quality of the experts' decisions. The second potential problem is the unavailability of experts in some geographical regions. The third risk is related to human error, which does not allow experts to achieve better results. In current methods of RMCDP, human error is an inevitable problem and it becomes more crucial when there is no controlling system for the experts' decisions [3]. The last and also main threat for RMCDP is the lack of automated processes. This can be critical when the demand for concrete, regardless of geographical location, is increasing throughout the world [6–12]. Current Portland cement production throughout the world will nearly double by 2050 [13]. In this paper, we introduce an automated RMCDP method at the tactical level by looking at this problem from a new angle. As has been mentioned, experts are handling RMCDP and we are attempting to match their decisions by using ensemble machine learning algorithms. Also, the size

* Corresponding author at: School of Civil and Environmental Engineering, Faculty of Engineering, The University of New South Wales (UNSW), Kensington, NSW 2052, Sydney, Australia. Tel.: +61 2 9385 6914, +61 450509295; fax: +61 2 9385 6139.
E-mail address: maghrebi@unsw.edu.au (M. Maghrebi).

http://dx.doi.org/10.1016/j.autcon.2016.03.007
0926-5805/

and amount of data that is used in this study is much greater than the datasets that have been used in similar research in the literature. The richness of the data helps the authors to draw their conclusions more confidently and introduce more generalized models. In this paper, first the context of RMCDP and the related literature are discussed. Second, in the Methodology section, the selected attributes are explained and the selected machine learning techniques are presented. In the third section, the features of the dataset acquired from the field are studied. Finally, the proposed idea is tested with field data and the results presented; the outcomes are discussed by comparing the selected machine learning techniques from a different perspective.

2. Related works

A text mining based study was performed on the body of literature that was devoted to "Ready Mixed Concrete" [14]. They showed that only in a few works the RMCDP were considered while concrete technology is the main core of research in this area. Moreover, as has been briefly stated, despite significant progress on RMC dispatching in the last two decades, many scholars have indicated the inefficiency of RMC dispatching and its dependence on human expertise [2,3,5,15]. A considerable amount of RMCDP literature has been published on mathematically modeling the RMCDP and solving the models heuristically. It has been proven that the RMCDP is an NP-hard problem [16–21]. This means that with available computing facilities we cannot solve large-scale RMCDP in polynomial time. A wide range of heuristic approaches have been implemented to address this issue and Genetic Algorithm (GA) has received more attention in the literature in comparison to other evolutionary methods [2,4,18,22–25]. Apart from GA, other methods include Particle Swarm Optimization (PSO) [26,27], Ant Colony Optimization (ACO) [28], Bee Colony Optimization (BCO) [29] and Tabu Search (TS) [29]. Although the different methods have been implemented, the discrete solution structure remains fairly much the same in most of the methods which consist of two merged parts: the first part defines the sources of deliveries and the second part expresses the priorities of customers. In the mentioned studies, the two most critical challenges are the number of infeasible allocations that exist in the initial solutions and computing time; this is because RMCDP has many side constraints that must be checked at each iteration, and it is also due to the random search behavior of the evolutionary methods. These methods mostly need a post-computing process to find feasible alternatives for infeasible allocations among the initial solutions. To overcome this issue Maghrebi, Waller and Sammut [18] presented an evolutionary based method which can solve the RMCDP without needing any additional algorithm and they developed a sequential meta-heuristic method which is 10 times faster than their previous method and rather than direct travel costs can also minimize the number of fleets [30]. More recently, Liu et al. [31] introduced an integrated framework for solving both production and delivery of RMC. Chou and Ongkowijoyo [32] present a decision aid model for selecting the on-site RMC type based on a reliability assessment process. Zhang et al. [33] integrated an intelligent monitoring system with a hybrid heuristic algorithm to more effectively reschedule RMCDP when customers' demands are assumed to be dynamic. Kinable et al. [34] introduced a new formulation similar to [16] but solved using constraint programming.

Beyond the heuristic approaches, other methods should be mentioned, such as Yan, Lai and Chen [21] who introduced a numerical method for solving the RMC optimization problem by cutting the solution space and incorporating the branch and bound technique and the linear programming method. Yan, Lin and Jiang [35] used decomposition and relaxation techniques coupled with a mathematical solver. Variable Neighborhood Search (VNS) was applied in RMCDP by Payr and Schmid [36]. More recently, Maghrebi et al. [17] implemented a Column Generation (CG) method which is amenable to Dantzig–Wolfe reformulation for solving large scale models which available computing facilities cannot optimally solve in polynomial time and this approach later on was compared with a heuristic method [37]. Similarly Benders decomposition was hired to near optimally solve RMCDP within a practical time [38]. Exploring experts' decisions in RMC dispatching centers was considered by [39].

The critical issue here is why the RMCDP is still being handled by experts. The reason behind the lack of success of the mathematically based models of RMC dispatching has been discussed in [3]. They found (i) the large number of variables and (ii) dealing with uncertain and dynamic data to be the two main obstacles in the models that attempt to solve the RMCDP optimally or near optimally. In this paper, we attempt to solve the RMCDP automatically. To this end, a wide range of machine learning techniques are used to match human decisions.

For these purposes, we intend to rely on experts' decisions in the RMC dispatching room and use their decisions for training the machine learning algorithm. A valid concern about this approach is the quality of the experts' decisions. Assessing the quality of experts' decisions is a cumbersome job. This is because: firstly, the quality of experts' decisions for large instances cannot be assessed due to the unavailability of the optimum solution for large instances; and secondly, RMC owners are pleased with the performances of experts because RMC dispatching jobs are still handled by experts who are able to find feasible solutions for day-to-day RMCDP [3]. This issue was extensively studied in [40,41] for relatively large instances (around 200 deliveries per day) by using high-performance computing facilities. They modeled RMCDP with soft time window (mixed-integer programming) and solid time window (integer programming) and tested these optimization models with different sizes of field data belonging to an active RMC. The optimization results were used as a benchmark and then compared with experts' decisions. The results show that in terms of cost, experts' decisions are around 90% accurate and are also more flexible in the event of an unexpected change in the system. They also argued that the experts' approach is totally different from the optimization models. In optimization, finding a feasible solution at the least cost is desirable, but in reality it is expected that experts supply all customers with the available resources and keep all the customers satisfied.

3. Methodology

In this paper we aim to introduce a method that can match experts' decisions automatically. In other words, we are looking for an alternative way of doing what is already done by experts in RMC dispatching rooms. To implement this idea, a wide range of supervised machine learning techniques are used. The training data includes the RMC monitoring data which covers all the information provided to the experts as well as the decisions that the experts have made. In particular, the dataset shows the experts' decisions in several circumstances. Therefore, it is expected that the selected machine learning techniques will match experts' decisions in any circumstances. In the attribute selection process, two issues have been taken into account: (i) the conducted research in this area such as [2,3,15,42–45]; and (ii) a consideration of the data that is already provided in practice to the experts which was determined after carefully observing the experts' behaviors in several RMC dispatching rooms. Then, the following parameters have been selected to construct the training and test datasets.

$$y = f(x) \tag{1}$$

y	experts' decisions about a selected depot for each delivery (decision variable)
f	machine learning technique (classifier)
x	input attributes set includes (parameters):
DOW	day of delivery in the week (Monday, Tuesday, …, Sunday)
VOD	volume of delivery (m^3)
EAT	expected arrival time at customer (hh:mm)
LON	longitude of customer

Please cite this article as: M. Maghrebi, et al., Matching experts' decisions in concrete delivery dispatching centers by ensemble learning algorithms: Tactical level, Automation in Construction (2016), http://dx.doi.org/10.1016/j.autcon.2016.03.007