Reinforcement Learning-based Large Neighborhood Search Approach to Dock Assignment and Truck Scheduling in Crossdocks

Shahin Gelareh[1,3], Roozbeh Sanaei[4], Rahimeh N Monemi[2,3]

[1] Département R&T, IUT de Béthune, Université d'Artois, F-62000 Béthune, France
shahin.gelareh@univ-artois.fr
[2] IESEG Lille, France
[3] Predictim Globe Ltd, Lille, France contact@predictim-globe.com
[4] SUTD - MIT International Design Centre, Singapore, Singapore

Mots-clés : *Logistics, Cross Docking, Assignment and Scheduling, Reinforcement Learning, Large Neighborhood Search*

1 Introduction

The question of whether and how combinatorial optimization algorithms can be automatically tailored to a avilable data for a family of problems has recently attracted significant interest from both machine learning and operations research communities [8, 2, 7].

A recently proposed resolution for undertaking these problems is to identify prime substructures through machine learning and thereafter leveraging a conventional solver as a generic black-box subroutine [1].

Reinforcement learning approach toward combinatorial optimization aims to substitute handcrafted heuristics of conventional algorithms with data-driven agents tuning their parameters according to received rewards and punishments based on their constructive or destructive conduct toward optimizing the provided objectives. However, this strategy is often challenging to apply in presence of intricate side constraints and cannot exploit the advantages of state-of-the-art commercial solvers. A recently proposed resolution for tackling this subject is to utilize large neighborhood search to leverage conventional solvers as a generic black-box subroutine. where the agent iteratively fixes a subset of variables and optimizes the rest through the black-box solver. In this contribution, we investigate how choice of deep neural architecture and reinforcement learning algorithm and their respective hyper-parameters affects the course of optimization in terms of convergence time and optimality of final solution, In particular, we consider three major algorithms of deep reinforcement learning, namely, deep Q-learning (DQN) and reinforce and actor-critic policy gradient methods as well as multi-layered and convolutional neural networks and study differences in behaviors of these algorithms in neighborhood selection for a cross-docking optimization problem. We focus on solving instances of dock assignment and truck scheduling problem arising within crossdocks. As this problem is solved on a regular basis at every such facility, we assume that the problem for every day has been generated from the same statistical distribution. Therefore, we use the historical data to learn from and we also expect that the unseen (future) observations will follow a similar distribution. The truck dock assignment problem, as many other optimization problems arising in cross dock management, has been receiving an increasing attention during the last decade. This is evidenced by several recent surveys ([6], [3], [9]) devoted to operations management and to scheduling in cross docks.

Our previous contributions on similar topics include [4] for analyzing different mathematical model and investigating their properties as well as a polyhedral and branch-and-cut approach in [5].

2 Problem Description

A set of trucks and a set of bi-function docks of a cross-dock are given. Every truck has an arrival time, a strict latest departure time, a docking time, which is the time spent for aligning the dock in front of a gate and the setups required by both truck and the dock to start the loading/unloading and a processing time which is required to load/unload the truck are given. In its general form, every truck belongs to different client and carries different cargo with different level of sensitivity translated to penalty cost, for every unit of waiting time before admission for service. A penalty cost is charged when a truck is not served at all. We would like to minimize the waiting cost of every single truck such that it can leave soonest possible and also to minimize the penalty weighted cost of missed trucks in the crossdocks day planning.

3 Our contribution

We implemented a compact while modular framework in which different constituting components of the algorithm can be exchanged handily to facilitate our experimentation. In terms of reinforcement learning strategt, we consider three major deep reinforcement learning strategies, namely, deep Q-learning (DQN) and REINFORCE and actor-critic policy gradient method and with regard to representation, mulit-layeed convolutional and multi-layered perceptrons are compared. PCA, TSNE, UMAP dimensionality reduction approaches. The truck dock assignment problem is stated in our newly proposed compact mixed integer programming formulation and solved as a black-box subroutine.

Références

[1] Maria-Florina Balcan, Travis Dick, Tuomas Sandholm, and Ellen Vitercik. Learning to branch. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 344–353. PMLR, 10–15 Jul 2018.

[2] Yoshua Bengio, Andrea Lodi, and Antoine Prouvost. Machine learning for combinatorial optimization : a methodological tour d'horizon, 2020.

[3] Nils Boysen and Malte Fliedner. Cross dock scheduling : Classification, literature review and research agenda. *Omega*, 38(6) :413 – 422, 2010.

[4] Shahin Gelareh, Fred Glover, Oualid Guemri, Saïd Hanafi, Placide Nduwayo, and Raca Todosijević. A comparative study of formulations for a cross-dock door assignment problem. *Omega*, 91 :102015, 2020.

[5] Shahin Gelareh, Rahimeh Neamatian Monemi, Frédéric Semet, and Gilles Goncalves. A branch-and-cut algorithm for the truck dock assignment problem with operational time constraints. *European Journal of Operational Research*, 249(3) :1144–1152, 2016.

[6] Jinxiang Gu, Marc Goetschalckx, and Leon F McGinnis. Research on warehouse operation : A comprehensive review. *European Journal of Operational Research*, 177(1) :1–21, 2007.

[7] Hao Lu, Xingwen Zhang, and Shuang Yang. A learning-based iterative method for solving vehicle routing problems. In *International Conference on Learning Representations*, 2020.

[8] Loris Serafino. *The No Free Lunch Theorem : What Are its Main Implications for the Optimization Practice ?*, pages 357–372. 01 2021.

[9] Jan Van Belle, Paul Valckenaers, and Dirk Cattrysse. Cross-docking : State of the art. *Omega*, 40(6) :827–846, 2012.