Delft University of Technology

Learning to Platforming

A Deep Reinforcement Learning Method for the Train Platforming and Rescheduling Problem

Zhang, Hongxiang; Zhu, Yongqiu; Hu, Liuyang; D'Ariano, Andrea; Wu, Yaoxin; Lu, Gongyuan

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Learning to Platforming: A Deep Reinforcement Learning Method for the Train Platforming and Rescheduling Problem

Hongxiang Zhang[1], Yongqiu Zhu[2], Liuyang Hu[1], Andrea D'Ariano[3], Yaoxin Wu[4], Gongyuan Lu[1]

[1]Southwest Jiaotong University, China, People's Republic of; [2]Department of Transport and Planning, Delft University of Technology, the Netherlands; [3]Department of Civil, Computer Science and Aeronautical Technologies Engineering, Roma Tre University, Italy; [4]Department of Industrial Engineering & Innovation Sciences, Eindhoven University of Technology, The Netherlands; hongxiang@my.swjtu.edu.cn

This paper proposes the Learning to Platforming (L2P) method, a novel graph neural network based deep reinforcement learning method, to solve the Train Platforming and Rescheduling Problem (TPRP). We customize a Markov decision process (MDP) to formulate the solving process of TPRP, utilizing a graph structure to represent states of trains, routes, and berthing tracks from a microscopic perspective. Then, we design a hybrid graph neural network named hAI-GNN to learn informative node embeddings on the graph encoded state. These embeddings are utilized to derive an effective action from the lightweight action space of MDP, which is associated with the decision object train under the state. A discrete-event simulation model is employed to serve as the environment of MDP and implement state transition mechanism. The hAI-GNN based policy network is trained by the Proximal Policy Optimization (PPO) algorithm with the reward function designed to minimize total knock-on delay trains and platform changes. The experiments on real-world instances show that the proposed L2P method can obtain high-quality solutions for both small and large scale instances within very short solving times.