

Reinforcement Learning Approaches for the TSP with Stochastic and Dynamic Release Dates

Yuanyuan Li, Claudia Archetti, Ivana Ljubic

IDS Department, ESSEC Business School, 95000, Cergy, France
{yuanyuan.li, archetti, ljubic}@essec.edu

Abstract. Traveling salesman problem with stochastic and dynamic release dates (RD) is a problem in which a company receives goods from its suppliers and distributes them to customers. Each customer is associated with a RD indicating the time when her parcel becomes available at the distribution center. The RDs are considered to be stochastic and dynamically updated during the distribution. We model the problem as a Markov decision process. We generate scenarios representing realizations of RDs and for each of them we approximate the future value using a batch approach. Two approximation approaches are proposed: policy function approximation through a consensus function in which a deterministic model determines the requests to serve immediately and those included in future routes, and a consensus function over all scenarios determines the final solution; one-step look-ahead policy with value function approximation where we build a two-stage stochastic model in which the first stage is to determine the route leaving immediately, while the second stage concerns future routes.

Keywords: Release date; TSP; Markov decision process