Lotte van Hezewijk

Multi-Echelon Inventory Optimization using Deep Reinforcement Learning

We study the applicability of a deep reinforcement learning approach to three different multi-echelon inventory systems, with the objective of minimizing the holding and backorder costs. We conduct an extensive literature review to map the current applications of reinforcement learning in multi-echelon inventory systems. We apply our deep reinforcement learning method to three cases with different network structures (linear, divergent and general structures). The linear and divergent cases are derived from literature, while the general structure case is based on a real-life manufacturer. We apply the Proximal Policy Optimization (PPO) algorithm, with a continuous action space. We show that the PPO solution outperforms the benchmark solution with 16.4% for the linear case, 8.3% for the divergent case, and 17.5% for the general case. Nevertheless, for large problem instances in the general case, the PPO algorithm is not stable, and we explain the limitations and avenues for future research.