

Применение технологий обучения с подкреплением для оптимизации управления запасами в цепочке поставок

Ахметжанов Р.Р. (Университет ИТМО),

Научный руководитель – к.ф.-м.н., доцент Ромакина О.М.

(Университет ИТМО)

Введение. Управление запасами в цепочке поставок (SCIM) является сложной задачей, требующей балансировки между потребностями клиентов и издержками. Эффективное управление запасами является ключевым элементом в обеспечении своевременного и качественного обслуживания клиентов, минимизации издержек хранения и предотвращении дефицита товаров. Традиционные методы управления запасами обеспечивают базовый уровень эффективности, но они часто недостаточны для решения сложных и динамически изменяющихся задач современных цепочек поставок. Современные методы машинного обучения, включая обучение с подкреплением (RL), позволяют находить оптимальные стратегии управления запасами в условиях неопределенности. В данном исследовании основное внимание уделяется применению алгоритмов глубокого обучения с подкреплением (DRL) в среде SCIM.

Основная часть. Методы обучения с подкреплением (RL) применяются для управления запасами в цепочке поставок, что позволяет находить оптимальные стратегии принятия решений. Традиционные методы, такие как Q-learning, сталкиваются с проблемами масштабируемости [1], в то время как глубокое обучение с подкреплением (DRL) использует нейросетевые модели, позволяющие работать с большими объемами данных [2]. В рамках исследования была разработана среда SCIM, включающая фабрику и распределительные склады, где агент принимает решения о производстве, хранении и транспортировке продукции. Формализация задачи проводилась в терминах марковского процесса принятия решений (MDP), включающего вектор состояния, вектор действий и функцию вознаграждения [3]. Эксперименты проводились с использованием методов оптимизации, таких как байесовская оптимизация и поиск по сетке, что позволило выявить параметры, обеспечивающие максимальный доход. Полученные результаты от этих экспериментов использовались, чтобы эффективно оценить, насколько успешно в той же среде может справиться агент DRL. Также в полученной среде был реализован один из алгоритмов DRL - DDPG, который продемонстрировал наилучшие результаты по эффективности управления запасами. Таким образом, исследование показало, что RL-алгоритмы способны значительно повысить эффективность управления цепочками поставок, снижая издержки и повышая адаптивность системы к изменениям спроса

Выводы. Данное исследование нацелено на исследование перспектив развития и потенциала методов RL в среде управления цепочками поставок. По результатам работ была получена эффективная среда для обучения DRL агентов в среде SCIM. В условиях полученной среды был разработан и применен алгоритм DDPG

Список использованных источников:

1. Chaharsooghi, S.K., Heydari, J., Zegordi, S.H.: A reinforcement learning model for supply chain ordering management: An application to the beer game. *Decision Support Systems* 45(4), 949–959 (2008).
2. Peng, Z., Zhang, Y., Feng, Y., Zhang, T., Wu, Z., Su, H.: Deep reinforcement learning approach for capacitated supply chain optimization under demand uncertainty. In: 2019 Chinese Automation Congress (CAC). IEEE (2019).
3. Yan, Y., Chow, A.H., Ho, C.P., Kuo, Y.H., Wu, Q., Ying, C.: Reinforcement learning for logistics and supply chain management: Methodologies, state of the art, and future opportunities. *Transportation Research Part E: Logistics and Transportation Review* 162, 102712 (2022).