DOCTORAL CONSORTIUM

Graph Neural Networks and Deep Reinforcement Learning for Adaptive Warehouse Optimization

DOI
https://doi.org/10.18690/um.fov.4.2026.77

ISBN
978-961-299-152-4

NEJC ČELIK, ANDREJ ŠKRABA
University of Maribor, Faculty of Organizational Sciences, Cybernetics & Decision Support Systems Laboratory, Kranj, Slovenia
nejc.celik1@um.si, andrej.skraba@um.si

Warehouse operations play a critical role in modern supply chains, with order picking representing one of the most resource-intensive processes in both manual and automated warehouses. Improving the efficiency of picking operations has therefore been a long-standing research focus in logistics and operations research. Traditionally, order-picking routing problems have been modeled as variants of the Traveling Salesman Problem (TSP) and addressed using heuristics and metaheuristics. Recent advances in machine learning, particularly deep reinforcement learning (DRL), offer promising opportunities for developing adaptive decision-making systems capable of operating in dynamic environments. At the same time, warehouse systems naturally exhibit graph structures that can be used by graph neural networks (GNNs). This doctoral research proposes an integrated framework combining GNNs and DRL to optimize warehouse processes.

Keywords:
order picking,
deep reinforcement learning,
graph neural networks,
warehouse routing optimization,
joint optimization

University of Maribor Press

1 Introduction

Warehouse efficiency has become increasingly important with the growth of e-commerce and the rising demand for rapid order fulfilment. Among warehouse operations, order picking typically accounts for the largest share of operational costs and labor time. As a result, improving the efficiency of order-picking processes has been a major topic in logistics and operations research. (Masae et al., 2020)

The order-picking routing problem was initially formulated as a variant of the Traveling Salesman Problem (TSP) (Ratliff & Rosenthal, 1983), focusing solely on minimizing travel distance between storage locations. More recent approaches extend this formulation to jointly address additional operational problems such as order batching, congestion management, and coordination among multiple pickers (Masae et al., 2020). A wide range of heuristics and optimization algorithms have been developed for this problem (Masae et al., 2020) and similar problems like vehicle routing problem (Psaraftis et al., 2016).

Advances in machine learning have enabled novel approaches to solving complex combinatorial optimization problems. Among these, deep reinforcement learning (DRL) provides a framework in which agents learn decision policies through interaction with an environment, making it particularly suitable for sequential decision-making problems such as routing (Mazyavkina et al., 2021). At the same time, warehouse systems can naturally be represented as graphs, where nodes correspond to entities and edges represent relationships between them. Graph neural networks (GNNs) have emerged as powerful models for learning from such graph-structured data. By aggregating information from neighboring nodes, GNNs can capture structural relationships within complex networks (Ju et al., 2024; Wu et al., 2021). In the context of warehouse systems, entities such as storage locations, pickers, items and orders can be represented as nodes and connected with edges in a graph, enabling the application of graph-based machine learning methods.

Despite these promising developments, the integration of GNNs and reinforcement learning in warehouse systems is still limited in existing research. This research therefore aims to investigate how GNNs combined with reinforcement learning can be used to develop adaptive decision-support systems for warehouse operations.

2 Research Problem

The central research problem of this doctoral work can therefore be summarized as follows:

How can warehouse systems be modeled as graphs and optimized using machine learning methods that combine graph neural networks and reinforcement learning?

The primary objective of this research is to develop a methodological framework for modeling warehouse systems as graphs and using this representation within a reinforcement learning framework to optimize order-picking operations.

More specifically, the research seeks to answer the following research questions:

RQ1: To what extent does the integration of graph neural networks and deep reinforcement learning improve routing decisions compared to traditional heuristic methods?

RQ2: To what extent can learned policies generalize or adapt to new problems and configurations in warehouse environments?

Addressing these questions will help determine whether graph-based machine learning approaches can provide a viable foundation for adaptive warehouse optimization systems.

3 Research Methodology

The proposed research will combine graph-based modeling, simulation, and machine learning methods to investigate adaptive optimization of warehouse processes.

First, warehouse layouts will be represented as graphs derived from relational data structures commonly used in warehouse management systems (WMS). In this representation, nodes will correspond to storage locations, items, orders, or other relevant entities, while edges will represent connections between them.

Second, graph neural networks will be used to process the graph representation of the warehouse and generate node embeddings that capture structural information about the environment. These embeddings will serve as inputs for the reinforcement learning agent, enabling the model to make decisions based on relational information within the graph.

Third, the order-picking problem, in combination with other warehouse operational problems, will be formulated as a sequential decision-making task suitable for reinforcement learning. In this framework, the agent represents a warehouse worker or robotic system performing warehouse operations. The objective of the reinforcement learning policy is to minimize the cost associated with operational actions while maximizing overall warehouse throughput.

Experiments will be conducted using simulated warehouse environments of varying sizes and configurations. The performance of the proposed approach will be evaluated by comparing learned policies with classical heuristics. Additional analysis will investigate the interpretability of learned embeddings and the ability of models to generalize across different warehouse layouts.

4 Expected Contributions

This research is expected to contribute to both logistics research and machine learning applications in operations management.

The research will propose a holistic GNN–DRL methodology for adaptive warehouse optimization. Additionally, this research will investigate the interpretability of learned graph representations and their potential use for decision support in warehouse management systems. Building on our work (Čelik & Škraba, 2026), this research will extend the application of GNN–DRL to more complex and realistic warehouse scenarios.

More broadly, the research aims to contribute to the development of adaptive, data-driven decision-support systems for internal logistics, bridging the gap between traditional optimization methods and modern machine learning approaches.

Acknowledgment

This research was funded by Slovenian Research and Innovation Agency (research core funding No. P5-0018), and the Ministry of Higher Education, Science, and Innovation of the Republic of Slovenia as part of the Next Generation EU National Recovery and Resilience Plan (Grant No. 3330-22-3515; NOO No: C3330-22-953012),

References

Čelik, N., & Škraba, A. (2026). Graph Neural Networks and Deep Reinforcement Learning for Warehouse Order Picking and Representation Learning. *Preprints*. https://doi.org/10.20944/preprints202602.0896.v1

Ju, W., Fang, Z., Gu, Y., Liu, Z., Long, Q., Qiao, Z., Qin, Y., Shen, J., Sun, F., Xiao, Z., Yang, J., Yuan, J., Zhao, Y., Wang, Y., Luo, X., & Zhang, M. (2024). A Comprehensive Survey on Deep Graph Representation Learning. *Neural Networks*, *173*, 106207. https://doi.org/10.1016/j.neunet.2024.106207

Masae, M., Glock, C. H., & Grosse, E. H. (2020). Order picker routing in warehouses: A systematic literature review. *International Journal of Production Economics*, *224*, 107564. https://doi.org/10.1016/j.ijpe.2019.107564

Mazyavkina, N., Sviridov, S., Ivanov, S., & Burnaev, E. (2021). Reinforcement learning for combinatorial optimization: A survey. *Computers & Operations Research*, *134*, 105400. https://doi.org/10.1016/j.cor.2021.105400

Psaraftis, H. N., Wen, M., & Kontovas, C. A. (2016). Dynamic vehicle routing problems: Three decades and counting. *Networks*, *67*(1), 3–31. https://doi.org/10.1002/net.21628

Ratliff, H. D., & Rosenthal, A. S. (1983). Order-Picking in a Rectangular Warehouse: A Solvable Case of the Traveling Salesman Problem. *Operations Research*, *31*(3), 507–521. https://doi.org/10.1287/opre.31.3.507

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(1), 4–24. https://doi.org/10.1109/TNNLS.2020.2978386