

# 폴필먼트 및 라스트마일을 위한 거점 물류센터의 DQN 강화학습 기반 최적 적재 알고리즘에 관한 연구

김승민, 하영호, 장현서, 송영은\*

고려대학교

songarak27@korea.ac.kr, haho9658@korea.ac.kr, kelvin926@korea.ac.kr, \*tdsong@korea.ac.kr

## A Study on the Optimal Loading Algorithm of the Local Logistics Center Based on Reinforcement Learning systems

Kim Seung Min, Ha Yeong Ho, Jang Hyun Seo, Song Young Eun\*

Korea University.

### 요약

본 논문은 코로나 시기 이후 증가한 물류량으로 인해 발생한 비용을 절감하기 위한 연구이다. 전 세계적으로 코로나 이후 물류량은 증가하고 있으며 그에 따라 지역 내 교통체증을 유발하고 라스트마일 배송을 위한 처리비용이 증가하고 있다. 이에 따라, 늘어난 물류 처리비용을 줄이기 위한 연구가 지속되고 있으며 본 논문에서는 이를 해결하기 위해 폴필먼트 및 라스트마일 배송을 처리할 수 있는 거점 물류센터 기반의 개념과 방식을 제시하고 성능 향상을 위한 강화학습 기반의 학습 테스트베드를 제시한다.

### I. 서론

최근 전 세계적으로 코로나 팬데믹 이후 전자상거래 및 온라인 쇼핑의 급증과 함께 물류량이 크게 증가하고 있다. 이러한 변화는 지역 내 교통체증을 야기하고, 라스트마일 배송을 위한 처리비용 증가를 초래하고 있다 [1]. 물류 산업은 이처럼 늘어난 물류량과 비용을 관리하기 위해 새로운 전략과 기술을 모색하고 있다.

특히, 라스트마일 배송은 물류체인의 최종 단계로, 제품을 고객에게 직접 전달하는 과정에서 가장 높은 비용과 복잡성을 발생시킨다. 이러한 문제점을 해결하기 위해 다양한 솔루션이 시도되고 있지만, 효과적인 방법을 찾는 것은 여전히 과제로 남아 있다.

본 논문에서는 증가한 물류량으로 인해 발생하는 비용 문제를 해결하기 위해 거점 물류센터 기반의 개념과 방식을 제시한다. 이 개념은 폴필먼트 및 라스트마일 배송을 처리할 수 있는 거점 물류센터를 중심으로 한 시스템을 구축하는 것이다. 이러한 접근법은 물류 프로세스를 효율화하고, 지역 내 교통체증을 완화시키며, 라스트마일 배송의 처리비용을 감소시키는 데 기여할 것으로 기대된다[2].

또한, 본 논문은 강화학습 기반의 학습 테스트베드를 통해 제안된 시스템의 성능 향상을 위한 학습 테스트베드를 제시한다. 이 테스트베드는 물류 프로세스의 효율성을 향상시키는 데 활용될 수 있는 최적의 전략을 찾기 위해 강화학습 기법을 적용한다. 이를 통해 거점 물류센터 기반의 개념이 어떻게 물류 비용을 절감하고 효율을 높일 수 있는지를 연구할 예정이다.

본 논문의 연구 결과는 코로나 시기 이후 증가한 물류량과 처리비용 문제를 해결하는 데 중요한 기여를 할 것으로 기대되며, 향후 물류 산업의 발전 방향을 제시할 것으로 예상된다.

### II. 본론

#### 1. 거점 물류센터의 개념 소개

그림 1과 같이 거점 물류센터는 기존의 대규모 물류센터와는 달리, 지역 내에서 일정한 범위를 처리할 수 있는 중소 규모의 물류 시설을 말한다. 이러한 거점 물류센터는 대규모의 물류 흐름을 지역 단위로 세분화하여 처리함으로써 효율을 높이고, 라스트마일 배송을 개선하는 역할을 한다.



그림 1. 폴필먼트 및 라스트마일 배송을 위한 거점 물류 센터 사진

#### 2. 거점 물류센터의 기대 효과

거점 물류센터의 운영 방식은 중앙 허브와 연계되어 운영되며, 거점 물류센터에서 물품을 수령하고 지역 내에서 배분 및 배송한다. 또한, 거점 물류센터는 지역 기반의 배송 네트워크를 구축함으로써 폴필먼트 처리비용 감소 및 라스트마일 배송 시간을 줄이고 비용을 절감할 수 있다. 효율적 분배 시스템을 통해 지역 내 주문을 효과적으로 분배 및 처리하여, 고객에게 빠르고 정확하게 배송할 수 있도록 한다. 또한, 지역 특성에 맞는 물류 운영을 통해 고객의 요구를 더 효과적으로 충족할 수 있다는 장점이 있다.

### 3. 강화학습을 위한 액션 및 보상 정의

Deep Q-Network(DQN)은 강화학습에서 사용하는 알고리즘 중 하나로, 심층 신경망을 활용하여 Q-러닝을 확장한 방법이다[3]. 본 논문에서는 강화학습을 이용하여 지속적인 학습과 최적화 과정을 통해 물품의 크기, 무게, 유형, 그리고 배송 경로 등 다양한 제약 조건을 충족시키며 최대한의 공간 활용을 이루어 낼 것으로 기대한다. 본 논문에서는 이를 위한 거점 물류센터의 액션 및 보상을 정의하고자 하였다.

표 1. 강화학습을 위한 거점 물류센터의 행동 정의

동작	구동 파츠	보상	동작 실패
0	컨베이어 이동 (좌)	1	이동 범위 초과
1	컨베이어 이동 (우)	1	이동 범위 초과
2	좌측 핸들러 이동 (앞)	3	물품 이중 입고
3	좌측 핸들러 이동 (뒤)	3	물품 이중 입고
4	우측 핸들러 이동 (앞)	3	물품 이중 입고
5	우측 핸들러 이동 (뒤)	3	물품 이중 입고
6	물품 입고	2	없음

### III. 결론

본 논문에서는 거점 물류센터의 개념과 방식을 고찰하고, 최적 물류 적재 알고리즘을 위해 강화학습을 적용하기 위한 거점 물류센터의 행동 정의를 제시하였다. 거점 물류센터는 지역 내 효율적인 물류 처리와 물품의 신속한 이동을 가능하게 하는 핵심적인 역할을 수행을 한다. 이러한 거점 물류센터의 운영 효율성과 효과성을 높이기 위해 강화학습 기법을 도입하여 최적의 물류 적재 경로와 작업 방식을 찾는 것이 필요하다.

이 논문에서는 구동부 파츠를 포함한 다양한 장비와 시스템을 최적화하기 위해, 물류센터의 작업 흐름과 목표를 고려한 강화학습 동작 구분 및 보상 체계를 제시하였다. 이러한 강화학습을 활용한 시스템은 물류 센터의 작업 효율, 안전, 에너지 효율성, 정확성을 향상시킬 수 있으며, 물류 처리 속도와 비용 절감에도 긍정적인 영향을 미칠 것으로 예상된다.

강화학습을 통한 물류 작업 최적화는 물류센터의 운영 방식을 혁신적으로 개선할 수 있는 잠재력을 가지고 있다. 지속적인 연구와 기술 개발을 통해 강화학습 기반의 물류 최적화 기법이 더욱 발전한다면, 물류 산업 전반의 경쟁력 향상과 지속 가능한 발전에 기여할 것으로 예상된다.

### ACKNOWLEDGMENT

본 과제(결과물)는 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다. (2021RIS-004)

### 참 고 문 헌

[1] Da Costa Barros, I.R., and Nascimento, T.P., "Robotic mobile fulfillment systems: A survey on recent developments and research opportunities," Robotics and Autonomous Systems Syst 137, 2021.

[2] Boysen, N., Fedtke, S. and Schwerdfeger, S. "Last-mile delivery concepts: a survey from an operational research perspective.," OR Spectrum 43, pp. 1 - 58 March 2021.

[3] Fan, J., Wang, Z., Xie, Y. & Yang, Z., "A Theoretical Analysis of Deep Q-Learning." Proceedings of the 2nd Conference on Learning for Dynamics and Control, 2020

[4] Winkelhaus, S., and Grosse, E. H., "Logistics 4.0: a systematic review towards a new logistics system." International Journal of Production Research, 58(1), pp. 18 - 43, 2020