

납기 위반 및 셋업 최소화를 위한 강화학습 기반의 설비 일정계획 모델

Machine Scheduling Models Based on Reinforcement Learning for Minimizing Due Date Violation and Setup Change

유우식(Woosik Yoo)*, 서주혁(Juhyeok Seo)**,
김다희(Dahee Kim)***, 김관호(Kwanho Kim)****

초 록

최근 제조업체들은 제품의 생산방식이 고도화 되고, 복잡해지면서 생산 장비를 효율적으로 사용하는데 어려움을 겪고 있다. 제조공정의 효율성을 방해하는 대표적인 요인들로는 작업물 종류 변경(job change)으로 인한 작업 준비 비용(Setup Cost) 등이 있다. 특히 반도체/LCD 공정과 같이 고가의 생산 장비를 사용하는 공정의 경우 장비의 효율적인 사용이 매우 중요한데, 상호 충돌하는 의사결정인 납기 준수를 최대화 하는 것과 작업물 종류 변경으로 인한 작업 준비 비용을 최소화 하는 것 사이에서 균형을 유지하는 것은 매우 어려운 일이다. 본 연구에서는 납기와 작업 준비 비용이 있는 병렬기계에서 강화학습을 활용하여 납기 및 셋업 비용의 최소화 목표를 달성하는 일정계획 모델을 개발하였다. 제안하는 모델은 DQN(Deep Q-Network) 일정계획 모델로 강화학습기반의 모델이다. 제안모델의 효율성을 측정하기 위해 DQN 모델과 기존에 개발하였던 심층 신경망 기반의 일정계획 생성기법과 휴리스틱 원칙의 결과를 비교하였다. 비교 결과 DQN 일정계획 생성기법이 심층신경망 방식과 휴리스틱 원칙에 비하여 납기 및 셋업 비용이 적은 것을 확인할 수 있었다.

ABSTRACT

Recently, manufacturers have been struggling to efficiently use production equipment as their production methods become more sophisticated and complex. Typical factors hindering the efficiency of the manufacturing process include setup cost due to job change. Especially, in the process of using expensive production equipment such as semiconductor / LCD process, efficient use of equipment is very important. Balancing the tradeoff between meeting the deadline and minimizing setup cost incurred by changes of work type is crucial planning task. In this study, we developed a scheduling model to achieve the goal of minimizing the due date and setup costs by using reinforcement learning in parallel machines with due date

This work was supported by Incheon National University Research Grant in 2018.

* Department of Industrial and Management Engineering, Incheon National University(wsyoo@inu.ac.kr)

** Department of Industrial and Management Engineering, Incheon National University(seojuhyuk@naver.com)

*** Department of Industrial and Management Engineering, Incheon National University(kmdj2003@nate.com)

**** Corresponding Author, Department of Industrial and Management Engineering, Incheon National University(khokim@inu.ac.kr)

Received: 2019-04-10, Review completed: 2019-08-09, Accepted: 2019-08-16

and work preparation costs. The proposed model is a Deep Q-Network (DQN) scheduling model and is a reinforcement learning-based model. To validate the effectiveness of our proposed model, we compared it against the heuristic model and DNN(deep neural network) based model. It was confirmed that our proposed DQN method causes less due date violation and setup costs than the benchmark methods.

키워드 : 일정계획, 강화학습, 납기, 셋업비용, 딥러닝

Scheduling, Reinforcement Learning, Due Date, Setup Cost, Deep Learning

1. 서 론

최근 제조업체들은 제품의 생산방식이 고도화 되고 복잡해지면서 생산능력을 효율적으로 사용하는데 어려움을 겪고 있다[2]. 제조공정의 효율성을 방해하는 대표적인 요인으로는 작업물 종류변경(job change)으로 인한 작업 준비비용(Setup Cost)이 있다[12]. 특히 반도체/LCD 공정과 같이 고가의 생산 장비를 사용하는 공정의 경우 장비의 효율적인 사용이 매우 중요하다. 따라서 효과적인 일정계획 수립시스템의 도입은 막대한 비용을 절감시킬 수 있는 최선의 방안이 되고 있다. 또한 효과적인 일정계획을 통해 고객과 약속한 납기를 최대한 반영시킨다면, 제조업체의 경쟁력을 크게 높일 수 있다[13]. 상호 충돌하는 의사결정인 납기 준수를 최대화 하는 것과 작업물 종류 변경으로 인한 작업 준비 비용을 최소화 하는 것 사이에서 균형을 유지하는 것은 매우 어려운 일이다.

예를 들어 반도체/LED Photo 공정의 경우 작업물 종류변경이 발생하게 되면 Mask를 교체해야하며, 타이어 가류 공정에서는 타이어 모양이 새겨진 몰드를 교체해야 하기 때문에 작업 준비비용이 발생한다. 작업물 종류 변경을 최소화 하는 것만 고려하여 일정계획을 생성한다면 생산 장비에 할당할 수 있는 작업이 제한되어 납기를 어기는 경우가 발생하여 고객

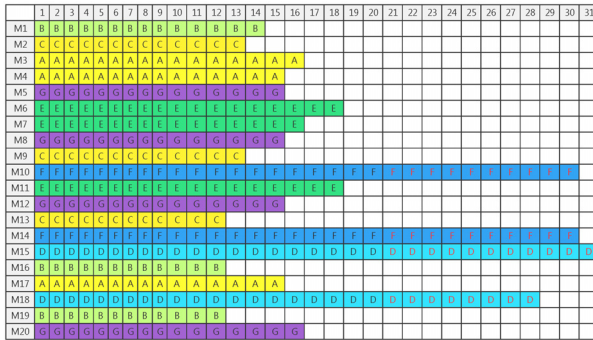
에게 신뢰를 잃는 경우가 생길 수 있다. 반대로 납기만을 고려하여 일정계획을 생성한다면, 작업물 종류 변경이 빈번하게 발생하게 되어 장비의 활용이 효율적이지 못하는 경우가 발생하게 된다. 따라서 작업물 종류 변경 횟수를 최소화하면서 납기 준수를 최대화 하는 일정계획을 수립해야 한다.

작업물 종류 변경이란 생산 장비에서 이전의 작업물과 현재 작업물의 종류가 달라지는 것으로 정의한다. 납기란 계약된 물품을 납품해야 할 납품 시기를 뜻하는 말로 본 연구에서는 제품이 생산 장비에서 작업이 끝나는 시점으로 정의한다.

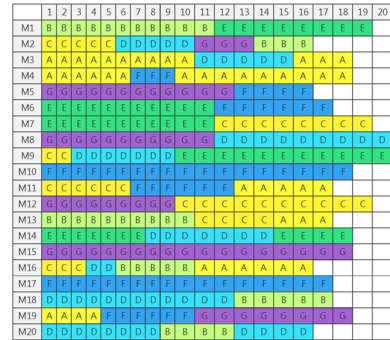
작업물 종류 변경만 고려한 일정계획의 예시<Figure 1>와 납기일 준수만을 고려한 일정계획 예시<Figure 2>는 다음과 같다.

작업물 종류만 고려한 일정계획에서는 작업 준비 비용이 발생하지는 않았지만, 납기일을 어긴 주문이 매우 많이 발생하였다. 납기일을 어긴 주문은 빨간색으로 표시하였다. 납기일 준수만 고려한 일정계획에서는 납기일을 어긴 주문은 발생하지 않았지만, 작업물 종류 변경은 빈번하게 발생하였다.

복잡한 스케줄링 문제를 해결하기 위해서 시뮬레이티드어닐링(Simulated Annealing), 유전알고리즘(Genetic Algorithm)과 같은 메타휴리스틱(Meta Heuristic)의 방법이 사용되어



〈Figure 1〉 Example of Machine Scheduling Considering Job Change Only



〈Figure 2〉 Example of Machine Scheduling Considering Setup Cost Only

왔다. 하지만 이러한 방법들은 현실의 복잡한 문제를 해결하는데 어려움이 있다. 메타 휴리스틱 기법은 다양한 해를 고려하지 못하며[8], 대규모 데이터셋과 탐색영역이 주어질 경우 성능을 보장할 수 없다는 문제점이 있다. 또한 해를 도출하기까지의 시간이 오래 걸린다는 단점이 있다[14]. 이러한 단점들로 인하여 제조 현장에서 적용하는 것에 어려움이 있었다.

이러한 문제점을 해결하기 위하여 머신러닝을 일정계획에 적용하는 다양한 연구들이 시도되고 있다. 머신러닝을 활용하여 일정계획을 생성하면, 복잡한 현업의 문제에도 적용이 가능하며, 메타 휴리스틱 기법에 비하여 시간이 절약된다.

강화학습을 활용한 일정계획에 대한 기존 연구로는 다음과 같다. 조용철은 인공 신경회로망(Artificial Neural Network)을 이용하여 반도체제조공정에서 웨이퍼의 이동경로의 최적 스케줄링 방법을 제안[1] 하였으며, Kim[9]은 DQN(Deep Q-Network) 알고리즘을 이용하여 무선 네트워크에서 DQN 패킷 스케줄링 알고리즘을 제안하였다[9]. Kim[7]은 강화학습을 이용하여 현재 스케줄링 상태를 고려하여 적합

한 휴리스틱 원칙을 선택하는 연구[7]를 하였다. Kang[5]은 Q-learning을 활용하여 단일기계에서 작업순서를 결정하는 의사결정 모형을 개발하였으며[5], Jeong et al.[4]은 강화학습을 이용하여 다양한 길이의 작업과 다양한 요구조건을 지닌 잡작 스케줄링 방법을 제안하였다[4]. 해외 연구로는 Gabel and Riedmiller[3]는 잡작 스케줄링 문제를 순차적 의사결정문제로 보고 독립적으로 할당을 결정하는 강화학습 모델을 개발하였으며[3], Shahrabia et al.[15]은 랜덤 작업 도착 및 기계 고정을 고려하는 스케줄링 방법을 강화학습을 이용하여 구현하였다[15]. 선행 연구에서 다양한 일정계획 분야에서 머신러닝을 응용한 기법이 제안되었고, 기존 휴리스틱 원칙 기반의 스케줄링 보다 개선된 결과를 확인하였다. 또한 머신러닝은 항생제 내성균 감염예측연구[10]나 장기간 태양광 발전량 예측연구[13] 등 다양한 예측연구 분야에 활용되고 있다.

기존 연구가 이동경로를 단축시키는 의사결정이나 가동률을 최대화하기 위한 의사결정, 최종 작업 완료시간(make span)을 최소화 하는

의사결정 이었다면, 본 연구는 납기 준수를 최대화하고, 작업물 변경을 최소화 시킬 수 있는 생산 장비에 작업을 배치하는 의사결정을 실시하는 것이므로 기존연구와 차이점을 가지고 있다.

본 논문에서는 강화학습(Reinforcement Learning)을 이용하여 병렬기계에서 제품과 주문량, 납기가 있는 주문의 리스트가 들어왔을 때, 납기 준수를 최대화 하고, 작업물 종류 변경을 최소화 하는 일정계획 문제를 해결하고자 하였다.

강화학습은 현재의 상태(State)에 따라 에이전트(Agent)가 미래의 기대되는 보상(Reward)을 최대화하기 위한 액션(Action)을 결정하는 인공지능의 한 방법이다. 이는 스케줄링 시 매 의사결정에 대한 가치는 명확하지 않지만, 최종적으로 주어진 보상 기준에 따라 더 좋은 스케줄 결과를 생성해야 하는 본 연구의 목표에 부합하는 방법론으로 판단된다.

강화학습을 이용하여 각각의 주문에 대한 상태를 생성하고, 그에 따른 보상으로는 작업물 종류변경의 횟수와 납기일을 어긴 주문의 횟수에 각각 가중치를 주었다. 행동으로는 Setup과 No Setup 두 가지 방식으로 하여 강화학습 모델을 만들었다. 본문에서는 제2장에서 제안기법을 설명하고 제3장에서 성능평가로 개발한 모델과 기존에 개발했던 심층 신경망(DNN)기반의 일정계획 모델, 휴리스틱 원칙(Heuristic Rule)의 결과와 비교하였다.

2. 제안기법

본 연구의 문제 상황은 단일공정에서, M개의 생산 장비가 있는 병렬기계에서 P개의 제품 종류가 있고, 제품 종류, 주문량, 납기일이 존재

하는 N개의 주문이 있다. 이러한 상황에서, 주문 $N_i(i = 1, \dots, N)$ 생산 장비 $M_j(j = 1, \dots, M)$ 에 할당할 때 납기를 준수하고 작업물 종류 변경을 최소화 하는 의사결정을 하는 것이다.

주문 유형에 대한 분포는 알려지지 않았으며, 각각의 작업은 순서에 독립적이라고 가정하였다. 각각의 주문은 여러 장비에서 동시에 작업 할 수 없고, 한 장비에서만 작업이 가능하다. 또한 장비 제약이 없이 모든 기계에서 모든 작업이 가능하며, 이때 장비 별 생산성도 동일한 것으로 가정하였다.

본 논문에서 사용하는 목적함수는 다음과 같다.

$$\text{Minimize } \sum_{i=1}^N \alpha \times \text{Job Change}_i \quad (1)$$

$$+ \beta \times \text{Due date violation}_i$$

α : penalty cost of job change

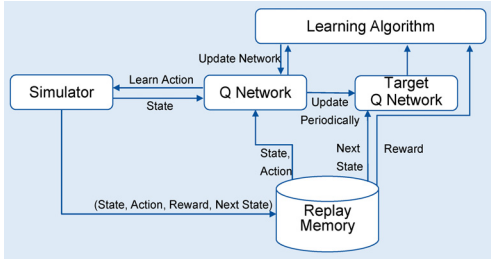
β : penalty cost of due date violation

$$\bullet \text{ Job change}_i : \begin{cases} 1, & \text{if a job change is} \\ & \text{occured by } N_i \\ 0, & \text{otherwise} \end{cases}$$

$$\bullet \text{ Due date violation}_i : \begin{cases} 1, & \text{if a due date violation} \\ & \text{is occured by } N_i \\ 0, & \text{otherwise} \end{cases}$$

강화학습을 이용하여 제품의 종류, 주문량, 납기가 있는 주문의 리스트에서 납기와 셋업 비용을 최소화하는 일정계획을 생성해주는 모델을 개발하였다.

<Figure 3>은 일정계획 모델의 구성도(Framework)이다. 시뮬레이터(Simulator)에서는 EDD-LPT순서로 정렬된 주문 리스트와 SLACK-EDD순서로 정렬된 주문에 대해서 상태(State), 행동(Action), 보상(Reward)을 만든다.



〈Figure 3〉 Framework of Proposed Model

첫째, 상태(State)는 다음과 같다. 일정계획을 생성할 때 상태로 고려될 수 있는 다양한 변수들이 있다. 이 중에서 Q-table의 상태 복잡도를 고려하여 의사결정에 도움이 되는 변수를 설정하여야 한다. 또한, 현재의 상황을 잘 고려하여 적절한 행동을 취할 수 있는 값이어야 한다. 본 연구에서는 장비의 현재 상태와 주문 정보, 과거의 납기 준수와 작업물 종류 변경 여부에 대해서 판단하여 상태를 생성한다. i 번째 주문에 대한 상태는 다음과 같이 정의된다.

$$s_i = [c_l^i (1 \leq \forall_i \leq P); p_l^i (1 \leq \forall_i p); \quad (2)$$

$$v^i; d^i; s^i; di^i]$$

- c_l^i : Number of machines setup for job type l
- p_l^i : Number of processed slots for job type l
- v^i : volume of order N_i
- d^i : due date of order N_i
- s^i : previous setup indicator
- di^i : previous due date violation indicator

c_l^i 는 제품 종류별로 현재 할당된 장비의 수를 나타내며, p_l^i 는 현재 할당된 작업의 수를 나

타낸다. v^i 는 현재 주문에 대한 주문량을 나타내며, d^i 는 현재 주문에 대한 납기의 정보이다. s^i 와 d^i 는 과거에 작업물 종류 변경의 발생 여부와 납기를 어기는 주문의 발생 여부를 판단한다. 판단 기준은 다음과 같다.

$$s^i = \begin{cases} 1, & \text{if the job type associated} \\ & \text{with the case order have casued} \\ & \text{a job change in the past} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

$$di^i = \begin{cases} 1, & \text{if the job type associated} \\ & \text{with the case order have casued} \\ & \text{a due date violation in the past} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

둘째, 행동(Action)은 다음과 같다. 일정계획 문제에 강화학습을 적용할 때 설정할 수 있는 행동에도 주문에 대한 생산 장비별 의사결정과 같은 다양한 경우가 있다. 하지만 행동의 수가 너무 많아지게 되면 복잡도가 증가하게 된다. 따라서 본 연구에서는 단순한 두 가지 방법으로 행동을 정의하였다. 상태 s_k 에 대한 행동은 다음과 같다.

- Action 1 : Without setup
 - Choose M^* which is not occur job change
 - Assign job k to machine M^*
- Action 2 : With setup
 - Choose M^*
 - Assign job k to machine M^*
 - M^* : smallest loaded machine

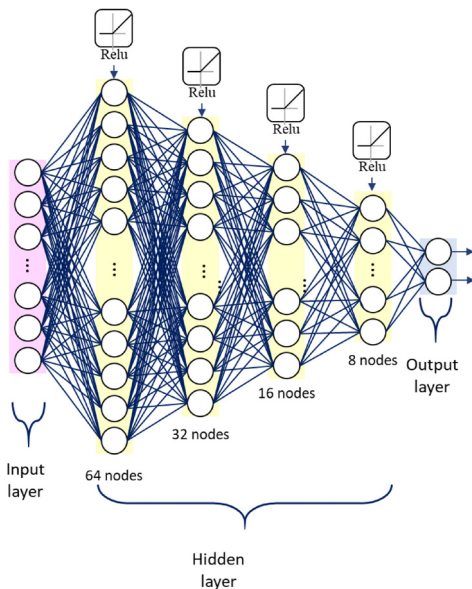
행동이 No setup일 경우 같은 작업물 종류 변경이 일어나지 않는 생산 장비 중 현재 할당된 작업이 가장 적은 곳에 작업을 할당한다.

그렇지 않을 경우에는 전체 생산 장비 중 할당된 작업이 가장 적은 곳에 작업을 할당한다.

마지막으로 보상(Reward)은 문제정의 부분에서 정의한 목적 함수와 같다(식 (1)).

상태, 행동, 보상이 만들어 지면, 만들어진 정보와 다음 상태(Next state)를 학습 데이터(Training data)에 저장하고, 현재 상태를 Q-Network에 저장한다. 저장된 학습데이터에서는 현재의 상태와 행동을 Q-Network에 전달하고, 다음 상태를 Target Q-Network에 전달한다. 학습 알고리즘(Optimizer)에서는 Q-Network의 상태, 행동과 학습 데이터의 보상을 가지고 손실(Loss)을 최소화 하는 학습을 하고, Q-Network를 업데이트 한다. Q-Network 부분에서는 주기적으로 Target Q-Network를 업데이트 하고, 학습의 결과를 시뮬레이터에 전달하게 된다.

학습에서 사용되는 Q-Network의 신경망 구조는 <Figure 4>과 같다. Q-Network는 1개의



<Figure 4> Structure of DQN Model

입력 층, 4개의 은닉 층, 1개의 출력 층으로 구성되어 있다. 모든 레이어는 완전 연결 층(Fully connected layer)으로 구성되어 있으며, 활성화 함수(Activation function)는 Relu(Rectified linear unit)를 사용하였다. 이 때 손실 함수(Loss function)으로는 평균 제곱 오차를 사용하였고, 학습 알고리즘(Optimizer)은 RMSProp optimizer를 사용하였다. 입력 값으로는 앞에서 생성한 상태가 들어오게 되고, 출력값으로는 행동에 따른 Q 값이 나오게 된다. 결과로 나온 Q 값 중 큰 값을 가지는 행동을 선택하게 된다. Target Q-Network의 구조도 Q-Network의 구조와 동일하다.

3. 성능평가

본 논문에서 제안한 강화학습 기반의 일정계획 생성 기법을 평가하기 위하여, 기존의 휴리스틱 원칙 기반의 일정계획 생성 기법의 결과를 비교하였다. 실험의 상황은 <Table 1>과 같다.

<Table 1> Experiment Environment

Warm up	200
Gamma	0.77
Epsilon	0.1
Target Q update frequency	200
Minibatch size	64
Buffer size	5000

비교대상으로 사용한 휴리스틱 원칙 기반의 생성 기법 종류는 총 3가지로 생성 기법은 다음과 같다[12].

우선 주문을 선택하는 방법은 다음과 같다.

- 1) EDD - LPT (Earliest Due Date - Longest Processing Time): 납기일이 짧은 순서대로 정렬 후 같은 납기일에 대해서 주문량이 많은 순서대로 재 정렬
- 2) EDD - MOR (Earliest Due Date - Most wOrk Remaining): 납기일이 짧은 순서대로 정렬 후 같은 납기일에 대해서 주문량이 많은 순서대로 재 정렬
- 3) SLACK - EDD: (납기일 - 주문량)의 값이 작은 순서대로 정렬 후 납기가 빠른 순서대로 정렬

다음으로, 설비를 선택하는 방법은 다음과 같이 단일규칙으로 하였다. 정렬된 리스트의 순서대로 주문을 작업물 종류 변경이 일어나지 않는 생산 장비 중 가장 작업량이 적은 곳에 할당한다. 만약 할당하였을 때 납기를 초과할 경우에는 전체 생산 장비 중 작업량이 가장 적은 곳에 할당한다.

이 때, 비용함수는 작업물 종류 변경 횟수 $\times 2$ (α), 납기를 준수하지 못한 주문의 수 $\times 10$ (β)으로 하였으며 총 비용은 이 둘의 합으로 하였다.

$$\sum_{i=1}^N = 2 \times Job\ Change_i + Due\ date\ Violation_i \quad (5)$$

$$Job\ change_i = \begin{cases} 1, & \text{if job change is occur on } i \\ 0, & \text{else} \end{cases}$$

$$Due\ date\ Violation_i = \begin{cases} 1, & \text{if Due date violation is occur on } i \\ 0, & \text{else} \end{cases}$$

스케줄링에서는 납기일을 준수하는 것이 셋업을 유발하는 것 보다 일반적으로 중요하므로, 현장전문가의 의견을 토대로 종합하여 설정하였는데 현장의 상황에 따라 다른 가중치가 필요한 경우가 존재한다.

3.1 실험 데이터 생성

제안 기법을 검증하기 위한 실험데이터의 생성 방법은 다음과 같다. 생산 장비가 M 개, 제품 종류가 P 개, 주문이 N 개의 주문에 대하여 최대 주문량($maxO$), 최소 주문량($minO$), 최대 납기일(K)을 입력 받는다. 다음으로 입력 받은 주문의 수만큼 제품 종류와 주문량, 납기일을 무작위(Random)로 생성한다. 주문량은 입력 받은 최대 주문량과 최소 주문량 사이의 무작위 값으로 생성하였으며, 납기일은 입력 받은 최대 납기일과, 최대 주문량을 처리할 수 있는 기간 사이의 무작위 값을 생성하였다. 이 때 항목별로 무작위로 생성되는 각각의 확률은 모두 동일하다.

본 논문에서는 생산 장비와 주문이 작은 상황과 생산 장비와 주문이 큰 상황에서 제품의 주문량이 같은 빈도로 발생할 때, 생산 장비가 한가한(Idle)일정 사례, 생산 장비가 적정한(Moderate)일정 사례, 생산 장비가 바쁜(Busy)일정 사례에 대해서 비교하였다.

3.2 설비 수와 주문량이 적은 사례

설비수와 주문량이 적은 사례는 생산 장비(M)가 16개, 제품종류(P)가 5개, 주문의 수(N)가 50개라고 하였다. 이 때, 생산 장비가 한가한(Idle) 사례, 생산 장비가 적정한(Moderate) 사례, 생산 장비가 바쁜(Busy) 사례에 대하여 10개씩 주문 리스트를 만들어 비교하였다.

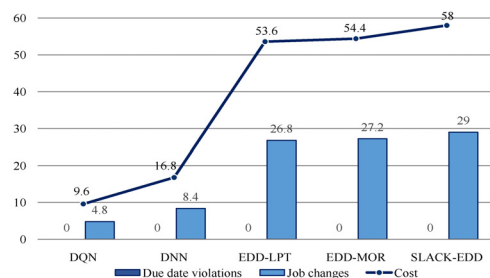
3.2.1 생산 장비가 한가한 일정 사례

생산 장비가 한가한 일정 사례는 최대 납기일을 15로 하여 주문 리스트를 만들었다. <Table 2>는 주문 리스트 10개에 대해서 작업물 종류

〈Table 2〉 Number of Job Changes and Due Date Violations in ($M=16$, $P=5$, $N=50$, $K=15$)

Scheduling Method Applied	DQN		DNN		EDD - LPT		EDD - MOR		SLACK - EDD	
Trials	#Job changes	#Due date violations	#Job changes	#Due date violations	#Job changes	#Due date violations	#Job changes	#Due date violations	#Job changes	#Due date violations
1	5	0	9	0	26	0	21	0	30	0
2	3	0	5	0	21	0	23	0	24	0
3	11	0	14	0	30	0	38	0	35	0
4	7	0	10	0	26	0	29	0	30	0
5	1	0	5	0	20	0	20	0	21	0
6	5	0	9	0	34	0	30	0	36	0
7	5	0	9	0	30	0	31	0	29	0
8	1	0	4	0	24	0	22	0	21	0
9	6	0	10	0	29	0	34	0	35	0
10	4	0	9	0	28	0	24	0	29	0
average	4.8	0	8.4	0	26.8	0	27.2	0	29	0
cost	9.6		16.8		53.6		54.4		58	

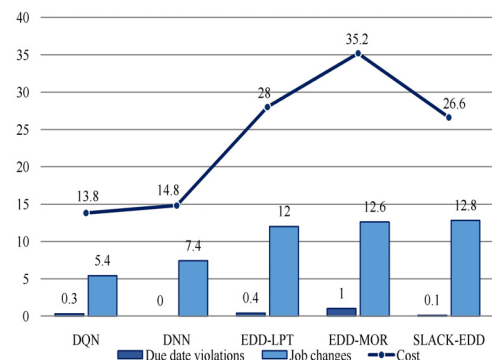
변경 횟수와 납기를 준수하지 못한 주문의 횟수를 나타낸 것이며, 〈Figure 5〉는 각각의 평균을 그래프로 나타낸 것이다.

〈Figure 5〉 Cost graphs of ($M=16$, $P=5$, $N=50$, $K=15$)

결과를 비교해보면 모두 납기일을 준수하였다. 하지만 작업물 종류 변경 횟수는 강화학습 기반의 일정계획이 다른 휴리스틱 원칙에 비해 훨씬 적은 것을 확인할 수 있으며, 심층 신경망 기반의 일정계획 생성기법 보다도 절반에 가까운 것을 확인할 수 있다.

3.2.2 생산 장비가 적정한 일정 사례

생산 장비가 적정한 일정 사례는 최대 납기일을 14로 하여 주문 리스트를 만들었다. 〈Table 3〉은 주문 리스트 10개에 대해서 작업물 종류 변경 횟수와 납기를 준수하지 못한 주문의 횟수를 나타낸 것이며, 〈Figure 6〉은 각각의 평균을 그래프로 나타낸 것이다.

〈Figure 6〉 Cost Graph of ($M=16$, $P=5$, $N=50$, $K=14$)

<Table 3> Number of Job Change and Due Date Violation in ($M=16$, $P=5$, $N=50$, $K=14$)

Scheduling Method Applied	DQN		DNN		EDD - LPT		EDD - MOR		SLACK - EDD	
Trials	#Job changes	#Due date violations	#Job changes	#Due date violations	#Job changes	#Due date violations	#Job changes	#Due date violations	#Job changes	#Due date violations
1	6	0	6	0	10	0	13	0	13	0
2	3	0	7	0	13	1	14	2	14	0
3	0	2	2	0	9	0	12	0	17	0
4	3	1	6	0	9	0	12	1	9	0
5	9	0	8	0	12	0	10	0	14	0
6	10	0	10	0	10	0	13	3	10	0
7	9	0	9	0	18	2	13	1	14	1
8	6	0	10	0	11	1	7	1	7	0
9	3	0	6	0	12	0	17	1	15	0
10	5	0	10	0	16	0	15	1	15	0
average	5.4	0.3	7.4	0	12	0.4	12.6	1	12.8	0.1
cost	13.8		14.8		28		35.2		26.6	

결과를 비교해보면 강화학습 기반의 일정계획이 심층 신경망 기반의 일정계획과 휴리스틱 원칙 중 SLACK 방법에 비하여 납기일을 어긴 경우는 많지만, 작업물 변경 횟수는 적은 것으로 나타났으며, 총 비용도 강화학습 기반의 일정계획이 더 적은 것을 확인하였다. 또한 휴리스틱 원칙에 비해서 총 비용이 절반 이하인 것을 확인하였다.

3.2.3 생산 장비가 바쁜 일정 사례

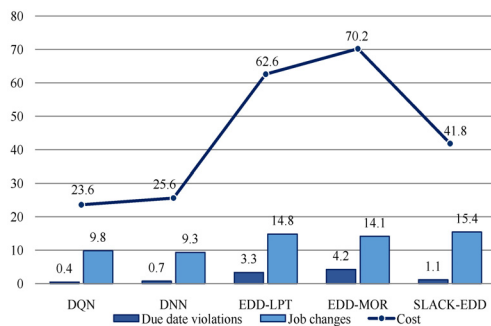
생산 장비가 바쁜 일정 사례는 최대 납기일을 13로 하여 주문 리스트를 만들었다. <Table 4>는 주문 리스트 10개에 대해서 작업물 종류 변경 횟수와 납기를 준수하지 못한 주문의 횟수를 나타낸 것이며, <Figure 7>은 각각의 평균을 그래프로 나타낸 것이다.

결과를 비교해보면 강화학습 기반의 일정계획이

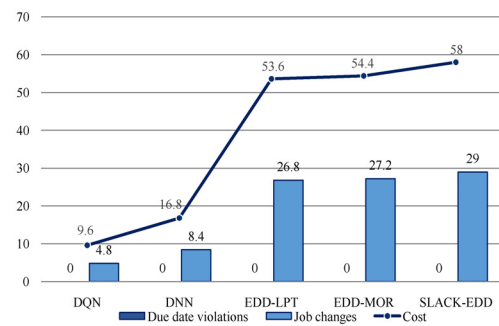
심층 신경망 기반의 일정계획에 비하여 납기일을 어긴 경우는 많지만, 작업물 변경 횟수는 적은 것으로 나타났으며, 총 비용도 강화학습 기반의 일정계획이 더 적은 것을 확인하였다. 또한 휴리스틱 원칙 중 LPT 방식이나 MOR 방식에 비해서 총 비용이 3배 가까이 차이나는 것을 확인하였다.

3.3 설비의 수와 주문량이 큰 사례

설비의 수와 주문량이 큰 사례는 생산 장비가 20개, 제품종류가 7개, 주문의 수가 100개라고 하였다. 이 때, 생산 장비가 한가한 (Idle) 사례, 생산 장비가 적정한 (Moderate) 사례, 생산 장비가 바쁜 (Busy) 사례에 대하여 10개씩 주문 리스트를 만들어 비교하였다.



〈Figure 7〉 Cost Graph of ($M=16$, $P=5$, $N=50$, $K=13$)



〈Figure 8〉 Cost Graph of ($M=20$, $P=7$, $N=100$, $K=24$)

〈Table 4〉 Number of Job Change and Due Date Violation in ($M=16$, $P=5$, $N=50$, $K=13$)

Scheduling Method Applied	DQN		DNN		EDD - LPT		EDD - MOR		SLACK - EDD	
Trials	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations
1	14	0	14	1	12	5	16	3	18	2
2	6	0	8	1	17	1	13	4	16	1
3	7	1	8	1	16	4	13	4	15	3
4	22	0	11	1	18	7	13	3	12	1
5	10	0	11	0	17	1	16	3	16	0
6	6	0	5	0	11	1	11	3	9	0
7	9	0	8	1	11	2	16	4	19	0
8	8	2	12	1	17	7	19	8	19	3
9	9	1	7	1	15	2	12	5	18	0
10	7	0	9	0	14	3	12	5	12	1
average	9.8	0.4	9.3	0.7	14.8	3.3	14.1	4.2	15.4	1.1
cost	23.6		25.6		62.6		70.2		41.8	

3.3.1 생산 장비가 한가한 일정 사례

생산 장비가 한가한 일정 사례는 최대 납기일을 24로 하여 주문 리스트를 만들었다. 〈Table 5〉는 주문 리스트 10개에 대해서 작업물 종류 변경 횟수와 납기를 준수하지 못한 주문의 횟수를 나타낸 것이며, 〈Figure 8〉은 각각의 평균을

그래프로 나타낸 것이다.

결과를 비교해보면 강화학습 기반의 일정계획만 납기일을 준수하지 못하는 경우가 발생하였다. 하지만 작업물 종류 변경 횟수가 다른 방식에 비하여 매우 작은 것을 확인할 수 있다. 따라서 총 비용도 강화학습 기반의 일정계획

<Table 5> Number of Job Change and Due Date Violation in (M=20, P=7, N=100, K=24)

Scheduling Method Applied	DQN		DNN		EDD - LPT		EDD - MOR		SLACK - EDD	
Trials	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations
1	5	0	9	0	26	0	21	0	30	0
2	3	0	5	0	21	0	23	0	24	0
3	11	0	14	0	30	0	38	0	35	0
4	7	0	10	0	26	0	29	0	30	0
5	1	0	5	0	20	0	20	0	21	0
6	5	0	9	0	34	0	30	0	36	0
7	5	0	9	0	30	0	31	0	29	0
8	1	0	4	0	24	0	22	0	21	0
9	6	0	10	0	29	0	34	0	35	0
10	4	0	9	0	28	0	24	0	29	0
average	4.8	0	8.4	0	26.8	0	27.2	0	29	0
cost	9.6		16.8		53.6		54.4		58	

기법이 작은 것을 확인하였다. 심층 신경망 기반의 일정계획 생성기법도 다른 휴리스틱 원칙에 비해서 작업물 종류 변경횟수가 작은 것을 확인하였다.

3.3.2 생산 장비가 적정한 일정 사례

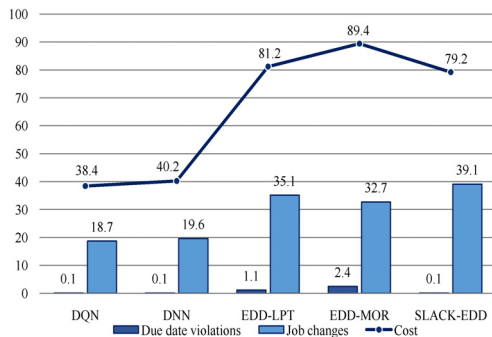
생산 장비가 적정한 일정 사례는 최대 납기일을 22로 하여 주문 리스트를 만들었다. <Table 6>은 주문 리스트 10개에 대해서 작업물 종류 변경 횟수와 납기를 준수하지 못한 주문의 횟수를 나타낸 것이며, <Figure 9>는 각각의 평균을 그래프로 나타낸 것이다.

결과를 비교해보면 강화학습 기반의 일정계획이 심층 신경망 기반의 일정계획과 SLACK

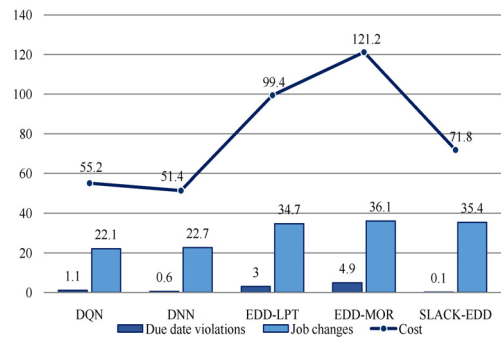
방법에 비하여 납기일을 어긴 경우는 많지만, 작업물 변경 횟수는 적은 것으로 나타났다. 하지만 총 비용은 심층 신경망 기반의 일정계획이 조금 더 좋은 것으로 나타났다. 휴리스틱 원칙에 비해서는 강화학습 기반의 일정계획 생성기법이 총 비용이 더 적을 확인할 수 있었다.

3.3.3 생산 장비가 바쁜 일정 사례

생산 장비가 바쁜 일정 사례는 최대 납기일을 20로 하여 주문 리스트를 만들었다. <Table 7>은 리스트 10개에 대해서 작업물 종류 변경 횟수와 납기를 준수하지 못한 주문의 횟수를 나타낸 것이며, <Figure 10>는 각각의 평균을 그래프로 나타낸 것이다.



〈Figure 9〉 Cost Graph of ($M=20$, $P=7$, $N=100$, $K=22$)



〈Figure 10〉 Cost Graph of ($M=20$, $P=7$, $N=100$, $K=20$)

〈Table 6〉 Number of Job Change and Due Date Violation in ($M=20$, $P=7$, $N=100$, $K=22$)

Scheduling Mathod Applied	DQN		DNN		EDD - LPT		EDD - MOR		SLACK - EDD	
Trials	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations
1	25	0	16	0	33	1	28	3	38	0
2	19	0	14	0	36	3	33	3	33	0
3	17	0	18	0	37	0	34	1	42	0
4	14	0	15	0	27	0	33	1	39	0
5	20	0	26	0	36	1	33	1	43	0
6	11	0	15	0	32	0	22	1	36	0
7	25	0	23	0	41	3	37	2	37	0
8	20	0	24	1	42	2	40	5	44	1
9	17	0	17	0	31	0	32	2	39	0
10	19	1	28	0	36	1	35	5	40	0
average	18.7	0.1	19.6	0.1	35.1	1.1	32.7	2.4	39.1	0.1
cost	38.4		40.2		81.2		89.4		79.2	

결과를 비교해보면 SLACK 방법이 납기일을 어긴 횟수가 가장 적은 것으로 나타났다. 또한, 강화학습 기반의 일정계획이 심층 신경망 기반의 일정계획에 비하여 납기일을 어긴 경우는 많지만 휴리스틱 원칙에 비해 적은 것으로

나타났다. 작업물 변경 횟수는 다른 방식에 비해 적은 것으로 나타났으며, 총 비용도 강화학습 기반의 일정계획이 더 적은 것을 확인하였다. 특히, MOR 방법에 비해서는 2배 가까이 차이는 것으로 확인되었다.

〈Table 7〉 Number of Job Change and Due Date Violation in ($M=20$, $P=7$, $N=100$, $K=20$)

Scheduling Method Applied	DQN		DNN		EDD - LPT		EDD - MOR		SLACK - EDD	
Trials	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations	#Job changes	#Duedate violations
1	24	2	24	0	36	4	39	5	27	0
2	21	1	23	0	30	1	34	3	34	0
3	22	1	22	2	31	4	30	6	38	0
4	31	0	30	1	36	2	36	5	39	0
5	16	2	17	1	32	3	36	4	28	0
6	25	0	21	0	34	5	40	8	34	1
7	24	1	22	1	34	4	33	5	35	0
8	17	0	19	0	35	2	36	1	35	0
9	14	3	24	0	36	2	37	5	42	0
10	27	1	25	1	43	3	40	7	42	0
average	22.1	1.1	22.7	0.6	34.7	3	36.1	4.9	35.4	0.1
cost	55.2		51.4		99.4		121.2		71.8	

4. 결 론

본 연구를 통해 강화학습을 이용하여 제품종류, 주문량, 납기가 정해진 주문 리스트에 대해서 납기와 셋업 비용을 최소화하는 일정계획 모델을 개발하고 기존에 개발한 심층 신경망 기반의 일정계획 생성기법과 휴리스틱 원칙 방식의 일정계획 방식과 비교하였다.

비교 결과 강화학습 기반의 일정계획 생성기법이 기존의 심층신경망 기반의 일정계획 생성기법과 휴리스틱 원칙 방식에 비하여 납기와 작업물 종류 변경 횟수가 적어 효율성이 높은 것으로 나타났다.

본 연구에서는 주문과 장비간의 제약이 없는 일정계획 문제에 대한 모델을 개발하였지만, 추후 장비 제약이 있는 문제나, 납기를 준수하기 위하여 주문을 분할하여 할당하는 문제등과 같이 복잡한 문제로 확장한다면 제조 현장의 문제에서도 적용할 수 있을 것이다.

References

- [1] Cho, Y. C., Cho, H. C., Kim, J. W., Zhang, L., and Jeon, H. T., "A Study on Optimal Scheduling of Multi-Spinner's Manufacturing Process Using Artificial Neural Network," International Journal of Fuzzy Logic and Intelligent Systems, Vol. 18, No. 1, pp. 157-160, 2008.
- [2] Demirkov, E. and Uzsoy, R., "Decomposition methods for reentrant flow shops with sequence-dependent setup times," Journal of Scheduling, Vol. 3, No. 3, pp. 155-177, 2010.
- [3] Gabel, T. and Riedmiller, M., "Distributed policy search reinforcement learning for job-shop scheduling tasks," International Journal of Production Research, Vol. 50,

- No. 1, pp. 41-61, 2012.
- [4] Jeong, H. S., Kim, M. W., Lee, B. J., Kim, K. T., and Youn, H. Y., "An Neural Network Approach to Job-shop Scheduling based on Reinforcement Learning," Proceedings of the Korean Society of Computer Information Conference, Vol. 26, No. 2, pp. 47-48, 2018.
- [5] Kang, J., "Feasibility Study on the Reinforcement Learning for a Dynamic Scheduling Problem," Journal of advanced engineering and technology, Vol. 10, No. 4, pp. 453-456, 2017.
- [6] Kang, Y. H., Lee, Y. S., and Shin, H. J., "Problem space based search algorithm for manufacturing process with rework probabilities affecting product quality and tardiness," Journal of the Korea Academia-Industrial Cooperation Society, Vol. 10, No. 7, pp. 1702-1710, 2009.
- [7] Kim, A. K., "Ensemble-based Quality Classification and Deep Reinforcement Learning-based Production Scheduling: Ensemble-based Quality Classification and Deep Reinforcement Learning-based Production Scheduling," Ph.D. dissertation, KyungHee University, 2018.
- [8] Kim, H. S., Park, D. J., and Jeong, I. J., "Railway Capacity Allocation Modeling using Genetic Algorithm," Transportation Research Record: Journal of the Transportation Research Board, Vol. 2608, pp. 1223-1230, 2016.
- [9] Kim, J. Y., "Packet scheduling algorithm using deep Q-Network in wireless network," MS dissertation, Yonsei University, 2018.
- [10] Lee, D. H. and Kim, K. H., "Deep Learning Based Prediction Method of Long-term Photovoltaic Power Generation Using Meteorological and Seasonal Information," The Journal of Society for e-Business Studies, Vol. 24, No. 1, pp. 1-16, 2019.
- [11] Lee, Y. S., "A study on Developing algorithms for Manufacturing networks using meta-Heuristics," Ph.D. Thesis, Sangmyung University, 2000.
- [12] Mazdeh, M. M., Nakhjavani, A. K., and Zareei, A., "Minimizing Total Weighted Tardiness with Drop Dead Dates in Single Machine Scheduling Problem," International Journal of Industrial Engineering & Production Research, Vol. 21, No. 2, pp. 89-95, 2010.
- [13] Oh, S. W., Lee, H. K., Shin, J. Y., and Lee, J. H., "Antibiotics-Resistant Bacteria Infection Prediction Based on Deep Learning," The Journal of Society for e-Business Studies, Vol. 24, No. 1, pp. 105-120, 2019.
- [14] Ryu, J. W., Kang, M. K., and Kim, M. W., "A Clustering Algorithm based on Heuristic Evolution Algorithm," Korean Institute of Information Scientists and Engineers, Vol. 27, No. 2, pp. 78-80, 2000.
- [15] Shahrabia, J., Adibib, M. A., and Mahootchia, M., "A reinforcement learning approach to parameter estimation in dynamic job shop scheduling," Computers and Industrial Engineering, Vol. 110, pp. 75-82, 2016.

저 자 소 개



유우식
1986년
1988년
1992년
1992년~현재
관심분야

(E-mail: wsyoo@inu.ac.kr)
서울대학교 공과대학 산업공학과 (학사)
한국과학기술원 산업공학과 (석사)
한국과학기술원 산업공학과 (박사)
인천대학교 산업경영공학과 교수
물류정보시스템, CAD/CAM, 생산정보시스템



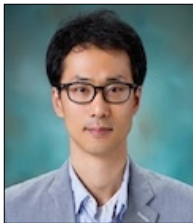
서주혁
2017년
2018년~현재
관심분야

(E-mail: seonjuhyuk@naver.com)
인천대학교 산업공학과 (학사)
인천대학교 산업경영공학과 석사과정
CAD/CAM, 생산정보시스템, 딥러닝



김다희
2016년
2017년~2019년
2019년~현재
관심분야

(E-mail: kmdj2003@nate.com)
인천대학교 산업공학과 (학사)
인천대학교 산업경영공학과 (석사)
㈜VMS solutions 사원
생산정보시스템, 인공지능



김관호
2006년
2012년
2013년
2014년~현재
관심분야

(E-mail: khokim@inu.ac.kr)
동국대학교 정보시스템전공 (학사)
서울대학교 산업공학과 (박사)
경희대학교 (연구박사)
인천대학교 산업경영공학과 교수
통계적 기계학습, 빅데이터, 딥러닝