

DQN기반 차량배차 자동화 모델

박승진, 마정목*
국방대학교 국방과학학과

DQN-based Vehicle Dispatch Automation Model

Seung-Jin Park, Jung-Mok Ma*
Department of Defense Science, Korea National Defense University

요약 우리나라의 인구 부족은 대표적인 사회적 문제이며, 그에 따른 대응 수단으로 자동화가 제시되고 있다. 자동화는 인공지능을 적용하여 더욱 확대되고 있으며 군에서도 감시정찰, 지휘통제 분야뿐 아니라 행정업무에 대해서도 자동화가 추진되고 있다. 본 연구에서는 군의 행정업무 분야 중 일일 단위로 반복되며 많은 시간이 소모되는 차량배차 업무에 대하여 다루었다. 군의 차량배차 업무는 배차신청이 접수되면 가용자원(차량, 운전병)을 편성해주는 것이다. 본 연구는 이러한 차량배차 업무를 대체 할 수 있는 차량배차 자동화 모델을 만들기 위하여 진행하였다. 연구에서 활용한 모델은 스케줄링 문제에 강점이 있는 강화학습을 선정하였으며, 그중에서도 기본 모델인 Q-learning을 개선하여 더욱 복잡한 환경에도 뛰어난 성능을 보이는 DQN(Deep Q-Network)을 활용하였다. DQN기반의 차량배차 자동화 모델은 가용 차량, 가용 운전병, 배차신청 내역을 바탕으로 차량배차 편성 수를 최대화하도록 모델링하고 학습을 진행하였다. 이후 모델의 성능이 사람을 대체할 수 있는 수준인지 판단하기 위해 사람과 모델을 대상으로 차량배차 편성실험을 진행하였으며, 그 결과 모델의 차량배차 편성량은 사람이 편성하는 것 대비 95.63%의 성능을 보였으며, 배차편성 시간과 편성오류 측면에서는 모델이 더 높은 성능을 보였다. 본 연구를 시작으로 향후 군 차량배차 업무의 자동화 및 효율성 증대가 이루어지길 기대한다.

Abstract The population shortage in Korea is a representative social problem, and automation is being proposed as a response measure. Automation is expanding further by applying artificial intelligence. Moreover, automation is being promoted in the military in surveillance, command and control, and administrative work. This study focused on the vehicle dispatch task, which is repeated daily and is time-consuming. The vehicle dispatch task of the military is to arrange available resources (vehicles, drivers) when a dispatch request is received. This study aimed to develop a vehicle dispatch automation model to replace this task. Reinforcement learning, which has strengths in scheduling problems, was selected as the model for this study. DQN, which showed excellent performance in a more complex environment by improving the basic model of Q-learning, was used. The DQN-based vehicle dispatch automation model was modeled and learned to maximize the number of vehicle dispatches based on available vehicles, drivers, and dispatch request records. Subsequently, to determine if the model performance could replace humans, a vehicle dispatch experiment was conducted with both human and model subjects. As a result, the vehicle dispatch volume of the model showed 95.63% performance of that of a person. The model showed higher performance in terms of dispatch time and dispatch error. Through this study, the automation and increased efficiency of vehicle dispatch tasks are anticipated.

Keywords : Reinforcement Learning, DQN, Automation, Vehicle Dispatch, Scheduling Optimization

*Corresponding Author : Jung-Mok Ma(Korea National Defense Univ.)

email: jxm1023@gmail.co

Received April 17, 2023

Accepted June 2, 2023

Revised May 24, 2023

Published June 30, 2023

1. 서론

2020년 기준으로 한국은 사망자 수가 출생자 수를 처음으로 넘어섰으며, 출산율은 0.84로 기록되었다[1]. 이와 같은 인구 문제의 대응 수단으로 자동화가 제시되고 있으며, 자동화는 제조, 물류, 금융 등 다양한 분야에서 적극적으로 활용되고 있다. 특히 자동화 기술은 스마트 물류센터, 스마트 공장과 같이 인공지능이 사용되면서 그 영역이 더 넓혀지는 추세이다.

인공지능을 적용하여 자동화하려는 노력은 국방 분야에서도 찾아볼 수 있다. 감시정찰 분야에서는 표적 식별 및 표적 정보 처리에 인공지능을 적용하였으며, 지휘통제 분야에서는 전장 정보를 자동으로 분석해주는 인공지능 참모에 대한 연구가 진행되고 있다[2,3]. 이와 같은 노력은 정보작전 분야뿐 아니라 행정업무에 대해서도 시도되고 있는데, 국방백서 2022에서는 인공지능 기술을 활용해 반복적인 행정업무를 자동화할 계획을 밝혔다[4]. 본 연구에서는 군의 다양한 행정업무 중에서 일일 단위로 반복되며, 많은 시간이 소모되는 차량배차 업무의 자동화를 다루고자 한다.

군의 차량배차 업무를 처리하는 과정은 다음과 같다. 먼저 간부들이 필요한 차량과 원하는 사용시간을 반영해 '배차신청'을 하고, 담당자는 접수된 배차신청에 대해 차량과 운전병을 편성한다. 하지만 가용자원(차량, 운전병) 대비 배차신청이 많은 경우 요청받은 신청 내역 전부를 반영하지 못하는 상황이 발생하며, 각각의 신청에 적합하지 않은 운전병 또는 차량을 편성하는 오류가 발생하기도 한다. 이러한 이유로 담당자는 차량과 운전병을 적절히 편성하여 요청된 배차신청 내역을 최대한 반영하고자 하며, 그와 동시에 발생 가능한 오류를 최소화하고자 한다. 이러한 차량배차 업무는 차량이라는 자원의 운용을 결정짓는 것으로 군에서뿐만 아니라 민간분야에서도 중요하며, 민간분야에서도 관련 연구들이 이루어졌다. 하지만 민간의 선행연구들은 물건을 효율적으로 운반하는 데에 집중되어 있으며, 배차편성 수를 극대화하려는 군에 바로 적용하기에는 한계가 있다. 그래서 군 차량배차업무의 특성에 맞는 모델이 필요하다.

군의 차량배차 업무는 두 가지 특징이 있다. 첫 번째로, 요청한 배차신청에 차량과 운전병을 동시에 편성하는 스케줄링 문제라는 것이다. 이는 앞서 언급한 민간분야 연구와의 차이에서 가장 두드러지는 점이다. 두 번째로 차량과 운전병을 편성할 때 고려해야 할 요소가 많아서 매우 복잡하다는 것이다. 이러한 두 가지 특징을 고려

했을 때 인공지능의 한 분야인 강화학습이 군의 차량배차 환경에 적합하다. 그 이유는 강화학습이 스케줄링과 같은 조합 최적화 문제에 강점이 있으며[5], 강화학습 기반의 알파고와 바둑을 학습한 것처럼, 강화학습은 복잡한 문제를 학습하는 데에 유리하기 때문이다.

따라서 본 연구에서는 강화학습을 군의 차량배차 환경에 적용하고자 하며, 강화학습의 모델 중 복잡한 환경에도 높은 성능을 보이는 DQN을 활용한다. 이를 통해 차량배차 업무에 대해서 사람을 대체 할 수 있는 DQN기반 차량배차 자동화 모델을 제시하고자 한다. 다음으로 모델의 성능이 사람을 대체할 수 있는 수준인지 판단하기 위해 사람과 모델을 대상으로 차량배차 편성실험을 진행하며, 차량배차 편성량, 소요시간, 정확성(오류 여부)을 비교하여 평가하고자 한다.

2장에서는 차량배차 관련 연구 및 강화학습 관련 연구를 설명하며, 3장에서는 DQN기반 차량배차 자동화 모델의 전반과 학습 과정을 설명한다. 4장에서는 성능 평가를 위해 사람과 비교한 실험과 그 결과를 설명한다. 마지막으로 5장에서는 본 연구의 결론 및 제한사항을 도출한다.

2. 선행연구

2.1 차량배차 관련 연구

차량배차에 관한 연구는 민간분야에서 주로 이루어졌으며, 많은 연구들이 휴리스틱 모형을 활용하여 수행하였다. 휴리스틱은 복잡한 문제를 단순화하고 단순화된 문제에 규칙을 적용하는 방법으로 문제를 해결하는 모형이다. 그 예로 위주영과 김병인은 쓰레기 수거차에 대한 배차문제에 휴리스틱 모형을 적용했다. 저자는 쓰레기 운반차의 경로를 만들고 그 경로에 가까운 고객을 최단 거리 경로로 추가하였다[6]. 이러한 규칙을 통해 총 이동 경로를 줄이고자 했다. 이명호도 휴리스틱 모형을 사용했다. 차량으로 컨테이너를 운반하는 문제에서 차량별 위치를 중심으로 컨테이너에 가까운 차량을 자동 배차하였다[7]. 앞서 제시한 연구와 달리, 김재광은 유전알고리즘을 사용했다. 유전알고리즘은 문제해결을 위해 여러 대안을 만들고 그중 가장 좋은 대안을 골라내는 방식이다. 저자는 물건 운반 작업에서 작업이 없는 차량뿐 아니라 작업이 종료예정인 차량도 다음 작업을 편성할 후보 차량으로 선정했다. 그렇게 더 많아진 후보 차량 중에서 하나를 선정해 작업을 편성하여 차량의 이용률을 높였다

[8]. 홍동희와 김귀정은 수학적 최적화 알고리즘을 활용하여 최적해를 구하는 솔버 모형을 사용했으며, 차량 대기시간과 공차 운행시간의 최소화를 목적으로 연구하였다[9].

2.2 강화학습 관련 연구

강화학습은 인공지능의 한 분야로, 주어진 환경에서 에이전트가 행동하고 그에 따른 보상을 받는 시행착오를 통해 학습한다. 그리고 최종적으로 보상을 최대화하는 행동을 찾아낸다. 본 연구에서는 이러한 강화학습을 차량배차 업무에 적용하고자 하며, 강화학습 모델로는 차량배차 시 고려요소가 복잡함을 반영하여 DQN을 선정하였다.

국방분야에서 DQN을 기반으로 한 연구는 무인무기체계와 군 통신체계에 대하여 이루어졌다. 먼저 고동현은 1대 1 교전 상황 간 무인전투기의 기동에 대한 연구에 DQN을 적용했다. 저자는 적항공기와 아군항공기의 속도와 위치에 대한 정보를 바탕으로 추격하거나 회피하는 행동을 하도록 모델링 하였다[10]. 다음으로 이동구 등 5명은 군용위성통신의 스펙트럼 효율을 높이기 위한 연구를 하였다. 스펙트럼 효율은 주어진 대역폭에서 전송되는 비트 수이며, 적응형 변복조는 무선통신 환경에 따라 변복조 기법을 가변적으로 할당하는 것이다. 저자는 SNR(Signal-to-Noise Ratio)의 값을 기준으로 변복조 기법을 변경하였으며, 변복조 기법이 달라지는 기준이 되는 SNR 값을 DQN모형을 통해 최적화하였다[11]. 앞서 제시된 두 가지 연구는 DQN모형이 국방 분야에 적용된 것이다.

민간분야에서의 연구 중 강형원 등 3명은 DQN을 이용해 반도체 공장의 장비에 작업을 편성하는 스케줄링 문제를 다루었다. 반도체 공장에는 다수의 장비가 있으며 주문이 들어오면 공장 내의 장비들에 작업을 편성한다. 이때 공장운영은 장비가 작업을 시작할 때부터 모든 작업이 완료되어 장비 사용이 종료될 때까지 이루어진다. 저자는 공장운영시간이 짧을수록 더 높은 보상을 주어 운영시간을 최소화할 수 있는 작업편성안을 만들어냈다[12]. 해당 연구는 장비에 작업을 편성하는 스케줄링 문제를 다룬다는 점에서 군의 차량배차와 공통점이 있다.

2.3 시사점

민간분야에서 이루어진 차량배차 관련 연구들은 이동 경로를 최소화하거나 차량의 이용률을 높이는 것을 목적

으로 연구되었다. 하지만 이러한 연구들은 접수된 배차 신청을 최대한 많이 편성하는 목적을 갖는 군의 배차업무와는 차이가 있으며, 선행연구를 군에 바로 적용하기에는 한계가 있다.

다음으로 강화학습에 대한 연구는 DQN모형을 기반으로 한 무인전투기 기동에 대한 연구와 군용위성통신의 스펙트럼 효율을 높이기 위한 연구들이 진행된 것을 볼 수 있었다. 하지만 군의 차량배차 업무에서와 같은 편성 최적화 문제를 다루지 않았다.

민간분야에서는 강화학습 DQN모형을 바탕으로 장비에 작업을 편성하는 편성 최적화 문제를 다루었다. 하지만 군의 차량배차업무는 작업이라는 1가지 요소를 장비에 편성하는 민간분야 연구와는 달리 차량과 운전병이라는 두 가지 요소를 배차신청에 편성해야 한다. 즉, 편성 최적화에 강화학습이 효과적임을 볼 수 있었으나, 배차신청에 차량과 운전병이라는 두 가지 요소를 편성할 수 있는 별도의 모델이 필요함을 알 수 있다.

3. DQN기반 차량배차 자동화 모델

3.1 이론적 배경

강화학습 모델 중 하나인 DQN은 주어진 환경에서 현재 '상태'와 '행동'을 입력으로 주면 다음 행동에 대한 기댓값을 출력해주는 Q 함수를 학습시킨다. 그리고 학습을 통해 최대의 보상을 받거나 최소의 벌칙을 받도록 에이전트의 행동을 결정한다[13]. DQN은 입력으로 현재 상태를 받으며 출력으로 취할 수 있는 행동에 대한 보상의 기댓값이 나온다. 그리고 그 기댓값을 바탕으로 Epsilon-Greedy 알고리즘에 의해 다음 행동을 선택한다. 학습은 Eq. (1)과 같은 비용함수를 최소화하는 방향으로 진행된다[14].

$$(r + \gamma \max_{a'} Q(s', a' | \bar{\theta}) - Q(s, a | \theta))^2 \quad (1)$$

Where, Q denotes action value function, s denotes State, a denotes Action, r denotes Reward, γ denotes damping factor, θ denotes neural net parameter vector

Eq. (1)에서 s 와 s' 는 각각 현재 상태와 다음 상태, a 와 a' 는 각각 현재 행동과 다음 상태에서 취할 수 있는 행동, r 은 현재 행동으로 인해 받은 보상이다. γ 는 감쇠

인자이며, 0과 1 사이의 값으로 설정하여 기대보상이 무한해지는 것을 방지하고 눈앞의 결과를 선호하는 것과 미래의 불확실성을 반영하도록 한다. θ 와 $\bar{\theta}$ 는 가중치로 이루어진 함수이고 각각 main network와 target network에 포함된다. 이렇게 main network와 target network를 나눔으로써 학습에 의해 가중치가 즉시 업데이트되기 때문에 발생하는 문제를 해결하였다. 또한 학습 간에 경험(상태, 행동, 보상, 다음 상태)에 대한 데이터를 저장하고 이를 무작위로 샘플링하여 학습에 재사용하는 Replay Memory 기법을 사용하여 학습 효율과 안정성을 높였다. 두 개의 상태가 서로 매우 비슷하지만 행동에 따른 보상이 다를 경우 학습이 잘 이루어지지 않는 파국적 망각 문제를 개선했다[15].

3.2 모델 정의 및 구성

본 연구에서 제안하는 DQN기반 차량배차 자동화 모델은 가용 차량, 가용 운전병, 배차신청내역을 바탕으로 차량배차를 편성해주는 모델로 Fig. 1과 같다. 본 모델은 다음과 같은 용어정리를 바탕으로 한다. 먼저 배차신청은 신청자의 요구를 담은 정보이며, "[신청자이름, 요구차종, 배차시작시간, 배차종료시간]"의 형식을 갖는다. 다음으로 배차신청내역은 날짜별로 종합된 배차신청을 모은 것으로 "[1번 배차신청, 2번 배차신청, . . .]"의 형식을 갖는다. 차량배차 편성은 배차신청내역에 있는 각각의 배차신청에 운전병과 차량을 편성하는 것이다. 차량배차 편성 수는 배차신청내역에서 차량배차 편성을

완료한 배차신청의 수이다.

DQN 기반 차량배차 자동화 모델의 구성은 Fig. 1을 통해 확인할 수 있으며, 다음과 같은 과정에 의해 학습이 이루어진다. 먼저 학습환경을 설정하고 데이터를 생성하는 과정이다. 여기에서 학습환경을 설정하는 것은 가용 차량과 가용 운전병에 대한 정보를 설정하는 것이며, 데이터 생성은 배차신청내역을 랜덤하게 생성해내는 것이다. 다음으로 가용 차량, 가용 운전병, 배차신청내역에 대한 정보를 0과 1로 이진 변환하여 만들어진 '상태'를 인공신경망에 입력한다. 그러면 배차신청 중 하나를 선택하는 '행동'을 한다. 그리고 그 '행동'을 바탕으로 배차 편성 및 보상지급 알고리즘을 거쳐 차량배차 편성 및 보상지급이 이루어진다. 다음으로 지급된 보상에 따라 모델의 파라미터가 업데이트되며, '상태'에 대한 정보가 업데이트돼서 다시 인공신경망에 입력된다. 모델의 학습은 이 과정을 반복하여 이루어진다.

3.3 모델 환경 설정 및 데이터 생성

모델의 기초가 되는 배차업무 환경(운전병 및 차량 정보)은 한 개의 보병 대대급 부대를 기준으로 했다. 다만 부대마다 차량 대수는 상이하므로 다르게 적용할 수 있다. 먼저 가용 운전병은 소형 운전병 2명, 중형 운전병 3명, 대형 운전병 1명, 총 6명으로 설정하며, 가용 차량의 종류는 총 7종으로 소형 4종('1t':1대, '2.5t':1대, 'Retona':3대, 'Korando':1대) 중형 2종('K511':5대, '5t':1대) 대형 1종('Midsize Bus':1대)으로 설정하여 연

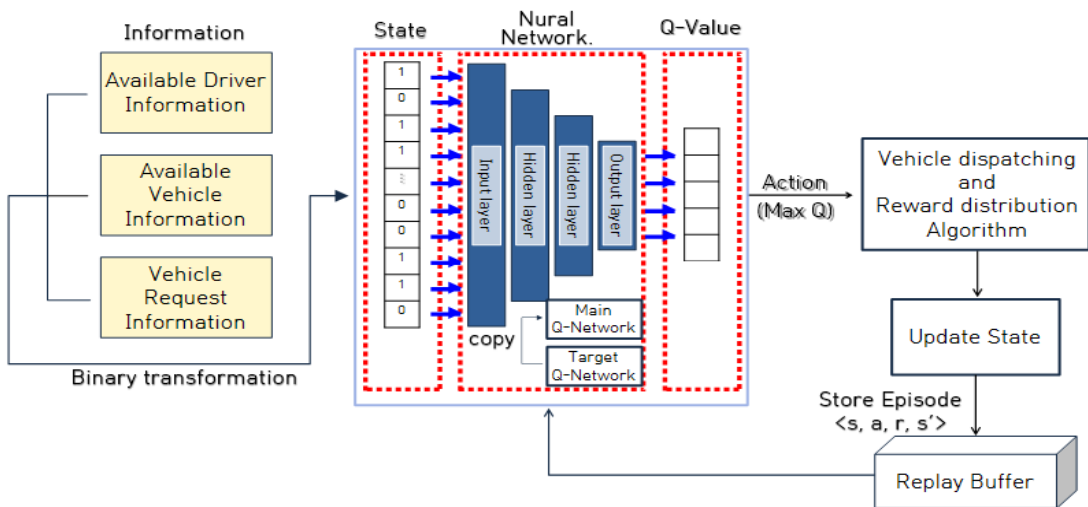


Fig. 1. DQN-based Vehicle Dispatch Automation Model

구를 진행한다.

본 모델의 학습데이터는 배차신청내역이다. 부대별 배차신청내역을 학습데이터로 활용해야 하나 군 보안상 실제 데이터를 사용하기 제한되어 랜덤생성하였다. 다만, 배차신청내역이 실제 데이터와 유사하게 생성될 수 있도록 데이터 생성 범위를 한정하였다.

배차신청내역은 배차신청이 모여서 생성되며, 배차신청은 "[신청자이름, 요구차종, 배차시작시간, 배차종료시간]"의 형식을 갖는다. 배차신청내역을 랜덤 생성할 때 신청자이름은 20개의 이름 리스트에서 임의 이름을 선택하도록 했다. 다음으로 요구차종은 7종의 차종 중 하나를 선택하도록 하였다. 다음으로 배차시작시간과 배차종료시간은 차량 이용시간과 출발 전 안전사항을 확인하는 시간을 고려해 최소 30분 이상 차이가 나도록 설정하였다. 그리고 해당 시간은 일과시간인 08시 30분부터 17시 30분 사이에 랜덤하게 설정되도록 한정하였다. 배차신청내역에 들어가는 배차신청의 수는 10에서 20 사이로 랜덤하게 설정하였다.

3.4 '상태' 구성

본 연구에서 '상태'는 0과 1로 이진 변환된 데이터로, Fig. 2과 같이 가용 운전병 정보, 가용 차량 정보, 배차신청정보로 구성된다. '상태'D를 형성하는 각 정보들의 구성은 Table 1과 같으며, 공통으로 포함된 것은 유효성(Validity)과 시간(Time)이다. 먼저 가용 운전병과 가용 차량의 유효성은 부상, 휴가, 고장, 정비 등으로 운행이 불가할 경우 0, 정상운행 가능할 경우 1로 표기된다. 그리고 배차신청정보에서 유효성은 배차신청이 존재하지 않거나 처리 불가하면 0, 그렇지 않을 경우는 1로 표기된다. 다음으로 시간은 총 48개의 0과 1로 구성된 정보로 24시간을 30분 단위로 나타낸다. 그래서 가용 운전병과 가용 차량의 시간에 대하여 30분 단위로 운전일정이 있으면 1, 배차가 편성되지 않아 운전일정이 없으면 0으로 표현된다. 배차신청정보에서는 각각의 배차신청에 있는 배차시작시간과 배차종료시간을 바탕으로 요구하는 시간에 대해서는 1, 나머지 시간을 0으로 표현한다. 다음으로 차종(Car type)은 총 7종류 차종에 대하여 나타낸 것으로, 가용 운전병 정보에서는 운행할 수 있는 차량을 1, 운행불가 차량을 0으로 표기하며 배차 신청정보는 요구하는 차종을 1, 요구하지 않는 차종을 0으로 표기한다. 이와 같은 방법으로 3가지의 정보가 표현되며, 예로서 배차신청정보가 Fig. 3와 같이 표현된 것을 볼 수 있다.

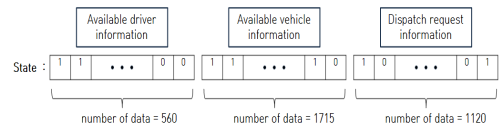


Fig. 2. Composition of State

Table 1. Composition of State in detail

Index	Composition		
Available driver information	Validity	Time	Car type
Available vehicle information			.
Dispatch request information			Car type

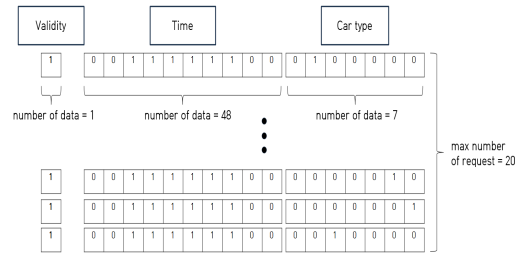


Fig. 3. Dispatch request information

3.5 배차 편성 및 보상 지급

3.4의 과정에 의해 형성된 '상태'는 인공지능망을 거쳐서 '행동'이 결정된다. 여기에서 '행동'은 1번에서 20번 배차신청 중 차량과 운전병을 편성할 하나의 배차신청을 선정하는 것이다. 다음으로 선택된 배차신청에 차량 및 운전병을 편성하는 것과 '행동'에 따른 보상을 지급하는 것은 Fig. 4의 '배차편성 및 보상지급 알고리즘'에 의해서 이루어진다.

해당 알고리즘은 크게 3가지 과정을 거친다. 먼저 정상적인 배차신청을 선택했는지 확인하는 과정으로 두 가지를 확인하게 된다. 첫 번째로 선택한 배차신청 번호가 생성된 배차신청 수 이하인지 확인한다. 예를 들어 15개의 배차신청이 있는 배차신청내역에서는 1번에서 15번 배차신청 중 하나를 선택했는지 확인하게 된다. 둘째, 선택한 배차신청이 아직 처리되지 않았는지 확인한다. 이 모든 조건을 순서대로 확인하여 충족하지 못할 경우 -1의 보상을 받고 종료한다. 다음 과정은 차량 및 운전병을 편성하는 과정이다. 먼저 차량은 배차신청에서 요구하는 차종에 대하여 첫 번째 차량부터 순서대로 확인하며, 이때 배차신청에서 요구하는 시간에 운행이 가능한지 확인

한다. 다음으로 운전병은 첫 번째 운전병부터 순서대로 확인하며, 이때 배차신청에서 요청한 차종을 운전할 수 있는지와 배차신청에서 요구하는 시간에 운행이 가능한지 확인한다. 이와같은 과정을 거쳐 운행 가능한 차량 및 운전병을 배차신청에 편성하고 보상을 +1 지급한다. 최종적으로 이와 같은 보상설정으로 정상적인 배차를 선택함과 동시에 최대한 많은 차량배차를 편성하도록 모델링하였다.

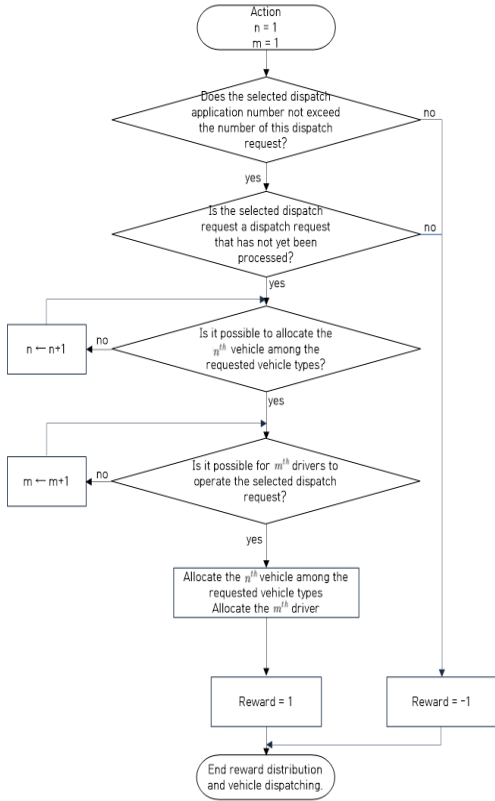


Fig. 4. Algorithm for vehicle dispatching and reward distribution.

3.6 '상태' 업데이트 및 학습결과

'행동'에 의해 선택된 배차신청에 차량과 운전병이 편성되며, 그에 따라 Table 1에 나타난 '상태'를 구성하는 가용 차량 정보, 가용 운전 정보, 배차신청정보의 값들이 업데이트된다. 첫 번째로 가용 차량 정보와 가용 운전병 정보의 시간 값이 변경된다. 총 48개로 이루어진 시간 값은 편성된 배차신청에서 요구한 시간에 해당하는 값이 0에서 1로 변경된다. 그리고 편성된 배차신청은 처리가 완료된 것으로, 배차신청정보에서 해당 배차신청의 모든 값이 0으로 변경된다. 또한, 편성된 운전병과 차량의 시

간 정보가 바뀔 때 따라 처리 불가능한 배차신청이 발생하며, 배차신청정보에서 처리 불가능한 배차신청에 해당하는 모든 값을 0으로 변경하여 '상태'를 업데이트한다. 그렇게 업데이트된 '상태'는 다시 인공신경망에 들어가며, 이러한 과정을 반복하여 모델의 학습이 진행된다. 최종적으로 지금까지의 과정을 바탕으로 학습을 진행하였으며, 학습 파라미터는 epoch는 10000, batch 크기는 200, 학습률 0.001, replay memory size는 1000으로 설정하였다. 그 결과 Fig. 5와 같이 Loss 값이 줄어들며 정상적으로 학습이 되는 모습을 볼 수 있다.

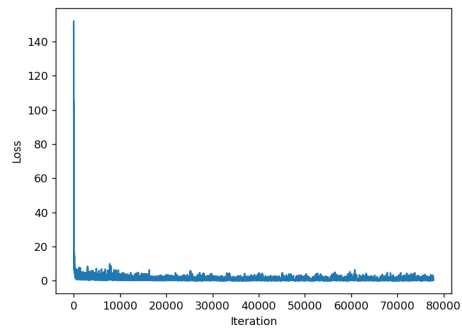


Fig. 5. Loss variation

4. 모델 실험 및 결과

DQN 기반 차량배차 자동화 모델이 사람을 대체할 수 있는지 평가하기 위해 실험을 진행하였다. 실험 대상은 학습된 차량배차 자동화 모델과 군의 차량배차 환경을 이해하고 있는 6년 차 이상의 육군 대위 5명으로 선정하였으며, 실험 대상에게는 가용 차량 정보, 가용 운전병 정보, 배차신청내역을 주었다. 그리고 이를 바탕으로 차량배차를 편성하도록 하였다. 실험은 배차신청내역을 변경하여 총 3회 진행되었으며, 실험 대상이 작성한 차량배차 편성에 대하여 차량배차 편성 수, 소요 시간, 편성 오류에 대하여 결과를 비교하였다.

실험의 결과는 Table 2와 같으며 모델의 결과와 사람 5명 결과의 평균을 비교하였다. 먼저 총 배차편성 수는 모델이 3번에 테스트에 대하여 13개, 11개, 11개이며 사람이 13.2개, 11.4개, 12개이다. 3번의 테스트의 평균을 비교했을 때 모델이 사람과 비교하여 95.63%의 배차편성 능력을 보였다. 다음으로 차량배차 편성을 하기까지 소요된 시간(Time)은 Model이 3번의 테스트 모두 1초 안에 수행했으며, 사람이 17분 32초, 11분 24초, 07

분 03초 소요됐다. 이를 통해 Model이 사람에 의해 편성하는 것보다 시간적 측면에서 우위에 있음을 볼 수 있었다. 마지막으로 편성오류(Error)는 사람의 경우 한 명의 운전병에게 운행시간이 중복되는 배차신청을 편성하는 중복배차 오류가 총 3건 발생하였으며 평균적으로 0.2회의 빈도를 보였다. 반면 모델의 경우 편성오류가 발생하지 않아 높은 신뢰도를 보였다. 결론적으로 차량배차 편성 수는 사람의 95.63% 수준이지만 소요시간이 적고 오류가 없다는 측면에서 사람을 대체 할 수 있는 수준의 결과를 보였다.

Table 2. Result of experiment

	Test 1		Test 2		Test 3	
	Model	Person(5)	Model	Person(5)	Model	Person(5)
The number of total dispatch (unit)	13	13.2	11	11.4	11	12
Time (MM:SS)	00:01	17:32	00:01	11:24	00:01	07:03
Error (unit)	0	0.4	0	0.2	0	0

5. 결론

본 연구에서는 차량배차 업무에 대해서 사람을 대체할 수 있는 DQN기반 차량배차 자동화 모델을 제안했다. 모델은 가용 차량, 가용 운전병, 배차신청내역을 바탕으로 '상태'를 구성하였으며, 이를 입력하여 하나의 배차신청을 선택하는 '행동'을 출력하도록 하였다. 그 다음 '배차편성 및 보상지급 알고리즘'을 거치고 '상태' 업데이트 및 '상태' 재입력의 과정을 통해 학습이 이루어지도록 하였다. 본 연구에서는 최종적으로 학습된 모델의 성능이 사람을 대체할 수 있는 수준인지 판단하기 위해 사람과 모델을 대상으로 차량배차 편성실험을 진행하였으며, 3가지 평가요소에 대하여 그 결과를 비교하였다.

실험 결과 본 모델은 사람과 비교하여 95.63%의 차량배차 편성 수를 보였으며, 소요시간과 편성오류 측면에서는 모델이 더욱 높은 성능을 보였다. 이를 통해 차량배차 편성 수는 사람 수준에 미치지 못하지만 효율성과 정확도 측면에서 사람을 대체 가능한 수준임을 확인하였다.

제한사항은 크게 3가지를 확인할 수 있었다. 첫 번째로 편성할 운전병과 차량을 선택하는 방식이다. 본 모델은 배차신청이 선택되면 해당 배차신청을 처리할 수 있

는 운전병 및 차량을 순서대로 확인 및 편성한다. 하지만 차량배차의 편성을 최대화 하기 위해서는 최적의 차량 및 운전병을 선택할 수 있는 별도의 모델이 필요로 하다. 두 번째로 본 모델은 시간을 30분 단위로 판단하도록 설정했으며, 이로 인해 하나의 배차신청의 배차종료시간과 다른 배차신청의 배차시작시간이 같은 30분 단위 안에 있으면 시간이 중첩된다고 판단하는 오류가 있었다. 세 번째로 다수의 신청자가 1대뿐인 중형버스를 요구할 때 일부 신청자의 차량을 5톤으로 조정하는 등의 유연한 판단은 불가능하다. 따라서 차후에는 운전병과 차량을 선택하는 부분에 대한 모델링과 설정한 시간단위에 따른 오류를 개선하며 추가적으로 대체 가능한 차량을 제안해 줄 수 있는 모델에 대한 연구가 요구된다.

References

- [1] S. J. Kim, 2020 Births 10% Decrease, Deaths 3.4% Increase...Population Decreased by 33,000, Emozak News, 2023, Available From: <http://www.emozak.co.kr/news/articleView.html?idxno=3177> (accessed April. 3, 2023)
- [2] J. K. Lee, J. D. Lee, J. K. Park, H. K. Yi, J. C. Ahn, S. H. Shin, H. Moon, S. M. Kim, B. J. Kim, "A Study on Research Trends and Future Direction of AI Models for Anomaly Analysis in Military Operations", *Journal of Digital Contents Society*, Vol.24, No.3, pp.631-640, 2023. DOI: <https://doi.org/10.9728/dcs.2023.24.3.631>
- [3] J. C. Shin, Instead of soldiers, AI watches the coast... North Korean wooden ship defection incident development, Culture Daily, 2021, Available From: <https://www.munhwa.com/news/view.html?no=20210927MW141423391984> (accessed April. 3, 2023)
- [4] Ministry of National Defense, 2022 Defense White Paper, Ministry of National Defense, Korea, pp.222-223, 2022.
- [5] What is Reinforcement Learning? 3 things you need to know, Mathworks, Available From: <https://kr.mathworks.com/discovery/reinforcement-learning.html> (accessed April. 3, 2023)
- [6] J. Y. Wi, B. I. Kim, "Heuristic method for rollon-rolloff vehicle routing problem", Proceedings of the Korean Industrial Engineering Society Spring Joint Academic Conference, Korean Institute Of Industrial Engineers, Korea, pp27-31, 2011.05.
- [7] M. H. Lee, Design of Heuristic Algorithm of Automatic Vehicle Delivery Support System, Digital convergence research, Vol.11, No.3, pp.181-187, 2013.
- [8] J. G. Kim, "Simultaneous Dispatching Policy for Transport

Vehicle in Container Terminal", Master's thesis, Pusan National University Graduate School, p.29.

- [9] D. H. Hong, G. J. Kim, "Optimizing the Vehicle Dispatching for Enhancing Operation Efficiency of Container Terminal", Journal of the Korean Convergence Society, Vol.8, No.10, pp.19-28, 2017.
- [10] D. H. Go, *Development of a Tactical Maneuver Algorithm for an Unmanned Combat Aircraft with Artificial Neural*, Ph.D dissertation, Inha University Graduate School, Incheon, Korea, pp.18-84, 2016.
- [11] D. G. Lee, Y. G. Sun, S. H. Kim, J. H. Sun, J. Y. Kim, "Novel Adaptive Modulation Technique based on Deep Reinforcement Learning for Military Satellite Communications" Novel Adaptive Modulation Technique based on Deep Reinforcement Learning for Military Satellite Communications", Proceedings of Symposium of the Korean Institute of communications and Information Sciences, Korea Institute Of Communication Sciences, pp.543-544, June 2022.
- [12] H. W. Kang, J. W. Lee, J. W. Huh, Scheduler based on the DQN that considers the following tasks to minimize tardiness"Scheduler based on the DQN that considers the following tasks to minimize tardiness", *Industrial Engineering Society of Korea Fall Conference symposium*, Korean, pp.3567-3580, November 2019.
- [13] R. S. Sutton, "Dyna, an integrated architecture for learning, planning, and reacting", ACM SIGART Bulletin, vol. 2, No.4, pp.160-163, Aug. 1991.
- [14] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, Marc G. Bellemare, A. Graves, M. Riedmiller, Andreas K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control though deep reinforcement learning," Nature, vol.518, pp.529-533, Feb. 2015.
- [15] S. E. Noh, Reinforcement learning from the ground up, p293, Youngjindotcom, 2020, pp. 56-204.

마 정 목(Jungmok Ma)

[정회원]



- 2002년 2월 : 육군사관학교 운영 분석 학사
- 2008년 8월 : 미국 펜실베이니아주립대(PSU) 산업공학 석사
- 2015년 5월 : 미국 일리노이대(UIUC) 산업공학 박사
- 2015년 9월 ~ 현재 : 국방대학교 국방과학학과 교수

<관심분야>

국방 모델링 및 데이터분석학, 무기체계 획득관리

박 승 진(Seung-Jin Park)

[정회원]



- 2016년 3월 : 육군사관학교 기계공학 학사
- 2022년 ~ 현재 : 국방대학교 국방과학학과 석사과정

<관심분야>

인공지능, 무인무기체계