

강화학습 기반 수검자 원내 이동경로 최적화

윤수영, 임소영, 유영준, 조환호*

*건양대학교 의료인공지능학과

sue2947@naver.com, imsy717@naver.com, ogoj20815@naver.com, *hhcho@konyang.ac.kr

Reinforcement Learning-Based Optimization of Intrahospital Movement Routes for Examinees

Suyeoung Yoon, Soyoung Im, Young Joon Yoo, Hwan-ho Cho*

*Department of Medical Artificial Intelligence, Konyang University

요 약

본 논문에서는 강화학습 기반의 수검자 원내 이동경로 최적화를 수행하였다. 최근 진단 시스템의 발전과 건강보험 급여대상의 확대로 의료 시스템 내 자원과 인력의 효율적인 배분은 의료 시스템 운영의 핵심 요소가 되었다. 물류 분배 시스템 분야에서는 목적지와 경로로 구성된 그래프에서의 경로 최적화 문제가 활발히 연구되어왔다. 본 연구에서는 원내 수검자 이동경로 최적화 문제를 해결하기 위하여, 이를 그래프에서의 이동거리 최적화 문제로 정의하였으며, 경로 최적화를 위해 Q-learning 을 적용하였다.

I. Introduction

최근 병원 진단시스템의 발전 및 건강보험 급여대상 항목 확장으로 인해 각종 진단 검사 시행 빈도가 의료인력 대비 증가하였다 [1]. 이에 따라 의료시스템 내의 제한적인 자원과 인력의 효율적인 배분은 의료 시스템 운영을 위한 필수적인 요소로 급부상하였다 [2]. 수검자 동선의 최적화는 원내 제한적인 인력과 자원을 효율적으로 활용할 수 있게 한다.

최근 물류 분배 시스템 분야에서는 배달 자원의 최적화를 위하여 강화학습을 적용하는 연구가 활발히 이루어지고 있다 [3, 4]. 물류 최적화는 일반적으로 목적지와 동선으로 구성된 그래프 내에서의 경로 최적화 문제로 접근한다. 원내 수검자 동선 최적화문제는 물류 분배 시스템의 배달 자원 최적화 문제 해결 방법과 유사하므로 비슷한 접근 방법을 이용해 해결이 가능하다.

따라서 본 연구에서는 수검자 동선 최적화 문제를 그래프 내 이동거리 최적화 문제로 정의하였고, Q-learning[5]을 이용한 이동경로 최적화를 적용하여 원내 수검자의 동선 최적화를 수행하였다.

가정하였고, 검사실 i 와 j 사이의 거리(D_{ij})는 검사실 i 에서 j 로 이동할 때 거쳐야 하는 격자의 개수로 정의하였다. 또한 각 검사실에서 소요되는 기본적인 검사 시간(분)을 reward 에 반영하기 위하여 다음 검사실 j 의 inherent cost (IC_j)를 정의하였다. 검사실 i 에서 j 로 이동할 때 최종적인 time cost TC_{ij} 는 아래 식 (1) 과 같이 정의되었다. D_{ij} 는 분단위의 IC_j 와의 scale 을 맞추어 주기 위해 D_{ij} 의 최대값을 이용해 normalization 을 수행하여 범위를 0~1로 제한하였다.

$$TC_{ij} = \frac{D_{ij}}{\max(D_{ij})} + IC_j \quad (1)$$

다음으로 검사실 i 에서 j 로 이동할 때의 최종 reward R_{ij} 를 아래의 식(2)을 이용하여 정의하였다. Reward 는 action 을 취했을 때 발생하는 time cost 와 음의 상관관계를 가지며, 최대값으로 normalization 하여 0~1 사이의 범위를 갖는다.

$$R_{ij} = 1 - \frac{TC_{ij}}{\max(TC_{ij})} \quad (2)$$

II. Methods

1) Problem description

문제의 정의를 위하여, 건양대학교병원 건강검진센터의 약도를 그래프로 변환하였다. 그림 1은 건강검진센터의 약도를 그래프로 변환하는 과정을 나타낸다. 수검자의 이동경로는 접수처에서 출발하여 모든 검사실(신체검사, 혈압측정, 시정력검사, 혈액채취, 소변검사, 흉부 X-ray 검사)을 방문한 후 다시 접수처에서 끝나는 것으로 가정하였다.

다음으로, Q-learning 에 사용할 reward 를 정의하기 위하여 검사실 간의 거리와 검사 소요시간을 이용한 time cost (TC_{ij})를 정의하였다(그림 2). 각 검사실은 높이와 너비가 균등한 격자 위에 위치해 있다고

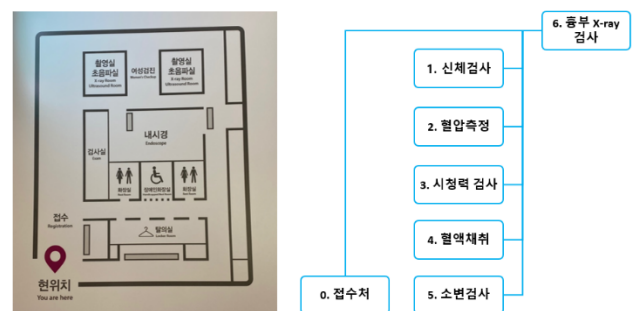


그림 1. (좌) 건강검진센터 약도, (우) 그래프 형태로 변환된 약도



그림 2. (좌) 격자 위에 설정한 각 검사실의 위치, (우) 각 검사실간의 이동 그래프, 그래프의 선의 굵기는 검사실간 거리의 역수를 나타냄

2) Q-learning

Q-table의 모든 파라미터는 100으로 초기화하였다. 이후 Q-table의 업데이트는 아래 식(3)을 이용하여 수행하였다 [5].

$$Q_{new}(S_t, a_t) \leftarrow Q_{old}(S_t, a_t) + \alpha \cdot (R + \gamma \cdot Q(S_{t+1}, \max_a) - Q_{old}(S_t, a_t)) \quad (3)$$

$Q(S_t, a_t)$ 는 Q-table의 각 요소를 의미하며, S_t 는 현재 상태로, [{지금까지 방문한 검사실의 리스트}, 마지막 검사실]으로 정의되었다. 또한 a_t 는 다음에 취할 action으로 다음 방문할 검사실을 의미한다. α 는 learning rate로 0.1, γ 는 discount factor로 0.9를 사용했으며, 이는 다음 action에서 기대되는 가장 큰 reward를 현재의 선택에 얼마나 반영할지를 의미한다. Q-table의 업데이트는 100,000번의 iteration을 통하여 이루어졌다.

III. Result

표 1은 디자인을 통해 구축된 reward matrix를 나타낸다. 표의 각 요소는 각 검사실간의 reward를 보여주며, 최대 reward는 1, 최소 reward는 0인 것을 확인할 수 있다.

표 1. Reward matrix

Room No.	0	1	2	3	4	5	6
0	0	0.37	0.70	0.0	0.5	0.66	0.21
1	0.89	0	0.81	0.11	0.61	0.76	0.29
2	0.87	0.46	0	0.12	0.62	0.78	0.28
3	0.86	0.45	0.81	0	0.64	0.79	0.26
4	0.84	0.44	0.79	0.12	0	0.81	0.25
5	0.83	0.42	0.78	0.11	0.64	0	0.24
6	0.9	0.46	0.79	0.09	0.53	0.75	0

그림 3은 epoch에 따른 최소 검사 소요 시간을 나타낸 것이다. 20000 epoch 이후에는 최소 검사 소요시간이 saturation되어 모델이 수렴한 것을 알 수 있다. 모델 최적화를 통해 얻은 검사를 가장 효율적으로

진행할 수 있는 최적의 이동 경로는 다음과 같다: [0. 접수처, 6 흉부 X-ray 검사, 1. 신체검사, 3. 시청력 검사, 4. 혈액채취, 5. 소변검사, 2. 혈압측정]. 이때 총 소요시간은 18.2분으로, 랜덤한 순서로 검사했을 때의 평균 소요시간 18.8분 (검사시간 16분 + [검사실간 평균 이동거리(9.5) × 7] / 이동거리 최댓값 24)에 비해 약 3.19% 시간을 절약할 수 있다.

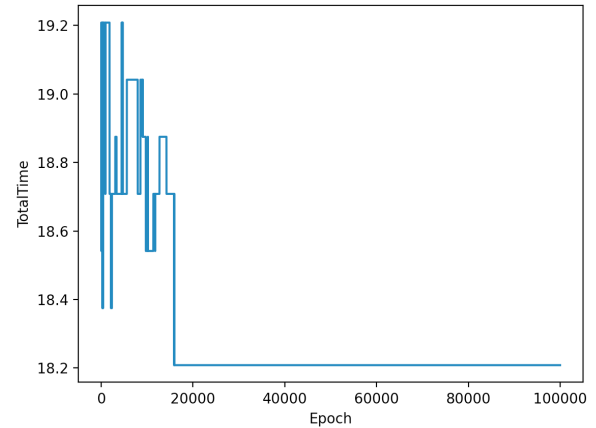


그림 3. Epoch 별 총 검사 소요 시간 (분)

IV. Discussion

본 논문에서는 강화학습을 이용하여 수검자의 원내 이동경로 최적화를 수행하였다. Q-learning을 적용하기 위하여 검사실간의 이동거리 및 검사 소요시간을 고려한 reward를 정의하였고, 이를 이용해 Q-matrix를 업데이트하여 최적의 경로를 도출하였다. 본 논문에서는 각 검사실의 대기시간은 고려되지 않았으므로, 실시간 수검자 배분을 위해서는 각 검사실의 실시간 대기시간을 반영할 수 있는 연구가 추가로 필요하다. 또한 검사실의 수가 6개로 매우 적은 한정적인 상황에 대한 시나리오를 이용했다는 한계가 있다. 추후 복잡도가 높은 동선 그래프에 대한 연구가 추가로 필요하다. 본 연구의 결과는 의료 시스템의 효율성을 증대하고 보다 나은 의료 서비스 제공에 기여할 수 있을 것으로 기대된다.

참 고 문 헌

- [1] Winder, Mateusz, et al. "Are We Overdoing It? Changes in Diagnostic Imaging Workload during the Years 2010–2020 including the Impact of the SARS-CoV-2 Pandemic." Healthcare. Vol. 9. No. 11. MDPI, 2021.
- [2] van Leeuwen, Kicky G., et al. "How does artificial intelligence in radiology improve efficiency and health outcomes?." Pediatric Radiology 52.11 (2022): 2087–2093.
- [3] Phiboonbanakit, Thananut, et al. "A Hybrid Reinforcement Learning-Based Model for the Vehicle Routing Problem in Transportation Logistics." IEEE Access 9 (2021): 163325–163347.
- [4] Yu, Lina, et al. "Reinforcement learning approach for resource allocation in humanitarian logistics." Expert Systems with Applications 173 (2021): 114663.
- [5] Watkins, Christopher JCH, and Peter Dayan. "Q-learning." Machine learning 8.3 (1992): 279–292.