현실 제약 조건을 반영한 다중 에이전트 강화학습 기반 다중 교차로 신호체계 최적화

Multi Intersection Traffic Signal Optimization based on Multi Agent Reinforcement Learning reflecting Real-World Constraints

박진혁[1] · 김혜민[2] · 전철민[3]
Jinhyuk Park[1] · Hyemin Kim[2] · Chulmin Jun[3]

[1]서울시립대학교 공간정보공학과 석사과정(jk0518@uos.ac.kr)
[2]서울시립대학교 공간정보공학과 석사과정(kimhm77@uos.ac.kr)
[3]서울시립대학교 공간정보공학과 교수 교신저자(cmjun@uos.ac.kr)

요 약

본 연구에서는 현실 제약 조건을 반영한 다중 에이전트 강화학습 기반 다중 교차로 신호체계 최적화를 진행하였다. 본 연구의 제안 모델과 단일 에이전트 강화학습 기반 모델, 현실 제약 조건을 제거한 비교 모델과의 비교를 진행하였다. 결과적으로 제안 모델은 단일 에이전트 강화학습 모델에 비해 대기 시간, 대기열 등 모든 면에서 더 좋은 결과를 보였다. 비교 모델에 비해 대기 시간, 대기열은 다소 떨어지는 결과를 보였지만 차량 정지 횟수가 크게 차이나는 것으로 나타났다. 본 연구의 제안 모델을 사용하면 실제 교차로에 큰 혼란 없이 적용할 수 있으면서 차량 대기 시간 및 탄소 배출량 등을 감소시킬 것으로 예상된다.

1. 서론

최근 기후변화 문제가 심각함에 따라 전 세계에서 기후변화 대응에 주목하고 있다. 기후변화 문제의 원인 중 하나로 지목되는 도로 혼잡 문제를 신호체계 최적화를 통해 완화하려는 연구가 활발하다. 최근에는 강화학습을 이용한 신호체계 최적화 연구가 등장하여 좋은 성과를 보이고 있다.

많은 강화학습 기반 신호체계 최적화 연구가 일반 강화학습을 사용하여 진행되었다. 하지만 실세계의 도로는 매우 복잡한 네트워크의 형태를 가지고 있어 도로의 교차로들은 인접한 교차로들의 영향을 받게 된다(김대호, 정옥란 2019). 이러한 특징을 일반 강화학습은 반영하지 못하기 때문에 다중 에이전트 강화학습을 이용한 신호체계 최적화 연구가 등장하였다. 다중 에이전트 강화학습을 이용한 신호체계 최적화 연구는 다양하게 존재하지만 모두 수학적 최적화에 집중하여 현실에 적용하기 어려운 모형을 제안하였다. 따라서 본 연구는 다중 에이전트 강화학습을 통해 교차로 간의 상호작용을 반영하고, 신호 순서 유지, 최소 녹색 시간 등 현실 제약 조건을 반영하여 현실성 높은 신호체계 최적화 모델을 제안한다.

2. 본론

본 연구의 시뮬레이션 환경은 실제 경기도 이천시의 연속된 6개의 교차로로 구성된다. 각 교차로들은 각기 다른 모양, 신호 주기, 도로 방향을 가지고 있으며 실세계 교통량을 반영하여 차가 많은 첨두 시간의 교통량을 사용하였다. 강화학습을 위한 상태, 행동, 보상은 다음과 같이 정의하였다.

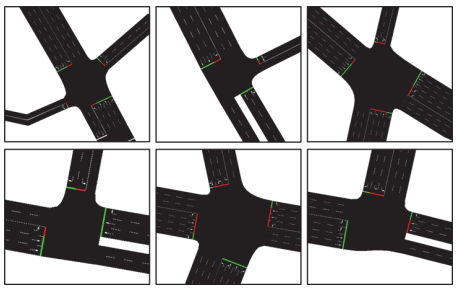

그림 1 시뮬레이션 교차로 환경

표 1 상태, 행동, 보상 정의

상태	[어느 신호가 켜져 있는 지, 현재 신호가 얼마나 켜져 있었는지(초), 신호 차선 별 차량 수, 신호 차선 별 차량 평균 속도]
행동	[0, 1], 고정된 신호 순서에서 0은 현재 신호를 유지, 1은 다음 신호로 변경을 의미한다.
보상	이전 step의 차량 대기시간 − 현재 step의 차량 대기시간

여기서 신호 제약 조건으로 기존의 신호 순서를 유지하도록 하였고 최소 녹색 시간과 최대 녹색 시간을 부여하여 너무 많거나 적은 시간을 한 신호에 할당하지 못하도록 하였다.

다중 교차로에서 인접한 교차로 간의 협력적 강화학습을 구현하기 위해서 기존 Q 함수 업데이트 식에 인접한 교차로의 이전 시간대의 보상을 더해주었다. 또한 교차로의 상태에 인접한 교차로의 상태, 행동을 추가해줌으로써 에이전트가 이웃의 상태와 행동을 고려하여 행동을 선택할 수 있도록 하였다.

실험은 본 연구의 제안모델과 다중 에이전트 강화학습의 효과를 비교하기 위한 단일 에이전트 강화학습 모델, 신호 제약 반영의 효과를 비교하기 위해 제안모델에서 신호 제약 조건을 제거한 모델을 학습하여 비교하는 것으로 진행하였다.

3. 결론

본 연구에서는 현실 제약 조건을 반영한 다중 에이전트 강화학습 기반 신호체계 최적화 모델을 학습하여 단일 강화학습 기반 신호 모델과 현실 제약 조건을 제거한 비교 모델과의 결과를 비교하였다. 결과는 차량 대기 시간, 차량 대기열, Co2 배출량 면에서 비교 모델, 제안 모델, 단일 모델 순으로 좋은 결과를 보였다. 차량 정지 횟수 면에서는 비교 모델이 제안 모델에 비해 크게 나타났다. 비교 모델은 특별한 제약 없이 신호를 바꾸어 대기시간을 가장 크게 줄였지만 신호를 자주 바꾸게 되어 차량들의 정지 횟수가 크게 나타난 것으로 판단된다.

“본 연구는 환경부「기후변화특성화대학원사업」의 지원으로 수행되었습니다.”

“This work is financially supported by Korea Ministry of Environment(MOE) as「Graduate School specialized in Climate Change」.”

참고문헌

1. 김대호, & 정옥란. (2019). 다중 교차로에서 협력적 교통신호제어에 대한 연구. 전기전자학회논문지, 23(4), 266-271.