

도시 교통 혼잡 완화를 위한 강화학습 기반 자율주행 차량 제어

오명진*, 황경호

국립한밭대학교 컴퓨터공학과

*omj0722@edu.hanbat.ac.kr, gabriel@hanbat.ac.kr

Self-driving vehicle control based on reinforcement learning to mitigate urban traffic congestion

Myung-Jin Oh*, Gyung-Ho Hwang

Dept. Computer Engineering, Hanbat National University

요약

본 논문에서는 급속한 자동차 보급과 함께 증가하는 교통혼잡비용 문제를 감소시키기 위해 작은 규모의 도시를 모델링한 도로에서 심층 강화학습 알고리즘인 PPO(Proximal Policy Optimization)를 적용한 자율주행 차량의 비율에 변화를 주어 전체 주행 차량의 속도에 미치는 영향을 조사하고 FLOW 시뮬레이터를 이용해서 성능을 검증한다.

I. 서론

인구 증가와 함께 차량 보유율이 높아지면서 도로 위 차량 수가 증가하고 도시 인프라의 부족으로 교통 체증이 증가하고 있다.

본 논문에서는 교통 혼잡 비용을 감소시키기 위한 대안 중 하나로 자율주행 차량의 도입을 제안한다.[1] 자율주행차량은 고도화된 센서와 인공지능 기술을 통해 효율적인 주행이 가능하며, 차량 흐름을 개선하고 도로 체증을 줄일 수 있다. 심층 강화학습 알고리즘 중 하나인 PPO를 적용하여 자율주행 차량의 도입이 도시 교통에 끼치는 영향에 대해 분석한다.[2]

자율주행 차량과 일반 차량의 비율이 교통 혼잡에 미치는 영향을 중점적으로 다루며, 도시 교통에 대한 새로운 전망과 해결책을 제시함으로써, 자율주행 기술의 발전이 도시 환경에서의 이동성과 편의성을 향상하는 데 어떤 영향을 미칠지에 대한 통찰력을 제공하고자 한다.

II. 본론

본 논문에서는 교통 시뮬레이터인 SUMO(Simulation of Urban Mobility)를 기반으로, 다양한 도로 네트워크와 차량 흐름을 시뮬레이션할 수 있는 FLOW와 강화학습 라이브러리인 RLlib을 접목하여 자율주행 차량의 성능 검증 시뮬레이션을 진행한다.[3][4] 시뮬레이션 환경에서 에이전트(Agent)는 자율주행 차량을 의미하며 주변 차량의 위치, 속도, 방향을 관측하여 에이전트가 판단을 내리는 데 필요한 상태(State) 정보를 받아 교통혼잡비용을 감소하는 최적의 속도를 학습하여 속도를 높이거나 줄이는 행동(Action), 또는 차선을 변경하는 행동을 수행한다. 도로의 최대 주행 속도는 $30m/s$ 로 설정하였으며 차량의 가속(accel), 감속(decel)의 최댓값은 $2m/s^2$ 로 설정하였다.

그림 1과 같이 시뮬레이션을 적용할 도로는 작은 규모의 도시를 모델링하여, 직선 도로, 곡선 도로, 교차로 등이 있으며, 1차선에서 4차선까지 다양하게 설계되었다.

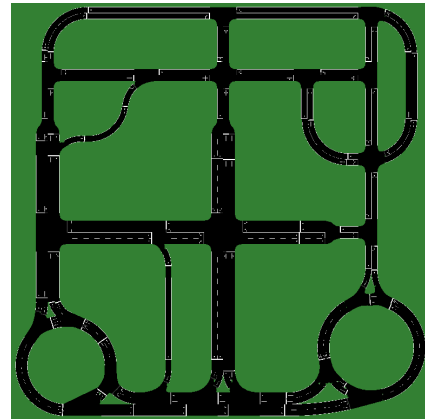


그림 1 MiniCity 도로

자율주행 차량은 IDM(Intelligent Driver Model)을 사용하여 인간 운전자의 가속을 모방한다.[5] 특히, 'obey_safe_speed' 모드로 설정하여 주변 환경 속도에 따라 안전한 속도에서 차로 변경이 가능하게 하였다. 비 자율주행 차량은 'no_lc_safe'(No Lane Change Safe) 모드로 설정하여 주변 환경의 속도에 관계없이 차로 변경이 제한되게 하였다. 차량들은 다양한 간격으로 배치되며, 최소 간격은 5m이다. 또한, 멀티 에이전트 환경을 사용하여 여러 에이전트가 동시에 환경과 상호작용하며 각자의 보상(Reward)을 최대화한다.

에피소드에서 에이전트가 취할 수 있는 최대 스텝 수는 3,000 단계로 설정하였고, 시뮬레이션 단계는 0.1초로 설정하여 300초 동안에 에피소드 1개가 진행된다. 이 시간 동안 자율주행 차량은 각자 보상을 최대화하는 행동을 결정한다.

$$\max(\|v_t\| - \|v - v_t\|, 0) / (\|v_t\| + \epsilon) \quad (1)$$

각 에이전트의 보상 수식은 수식 (1)과 같다. 보상은 시스템에 존재하는 모든

차량의 속도를 기반으로 설정되었으며, v 는 차량의 현재 속도 벡터, v_t 는 목표 속도(20m/s)를 나타내는 벡터, 그리고 ϵ 는 0으로 나누는 것을 방지하기 위한 아주 작은 양수이다. 차량의 현재 속도와 목표 속도 사이의 차이가 작을수록, 즉 차량이 목표 속도에 가까이 있을수록 보상이 커지게 된다. 보상은 0 이상의 값을 가지며, 차량의 속도가 목표 속도보다 빠를 때 음수의 보상을 반환하는 것을 방지한다.

심층 강화학습 알고리즘인 PPO에서 사용되는 모델은 완전 연결층이 3개 있고, 각 층은 32개의 노드를 가지고 있다. 훈련 배치 크기는 3,000으로 설정하였으며, 할인 계수(γ)는 0.999로 설정하고 GAE(Generalized Advantage Estimation) 램다(λ)는 0.97로 설정하였다.[6] 각 훈련 단계에서 수행하는 SGD(Stochastic Gradient Descent)의 반복 횟수는 15로 설정하여 시뮬레이션을 진행하였다.[7]

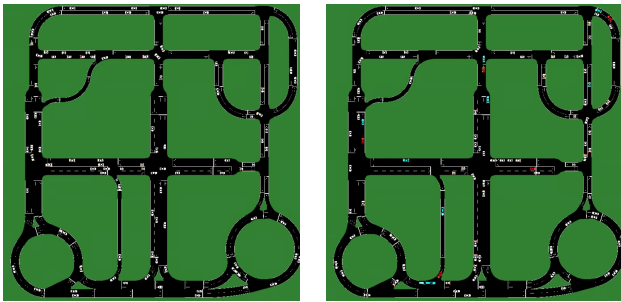


그림2 MiniCity - 자율주행 차량 비율 0%(좌), 5%(우)

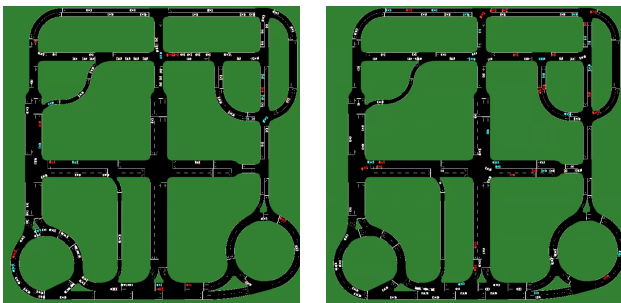


그림3 MiniCity - 자율주행 차량 비율 10%(좌), 20%(우)

자율주행 차량 비율	0%	5%	10%	20%
전체 차량 평균 속도(m/s)	4.78	7.85	7.62	7.39

표 1 자율주행 차량 비율별 평균 속도

그림 2와 그림 3은 MiniCity 네트워크에서 자율주행 차량의 비율을 변경하여 진행한 시뮬레이션 결과를 보여준다. 100대의 차량이 도입되었으며, 순서대로 0%(비 자율주행), 5%, 10%, 20%의 자율주행 차량의 비율에 변화를 주어 Mini City 환경에 배치되었다. 총 5,000 에피소드 동안 학습이 이루어졌으며, 표1과 같이 전체 차량의 평균 속도는 자율주행 차량 비율이 0%인 경우 4.78m/s, 5%인 경우 7.85m/s, 10%인 경우 7.62m/s, 20%인 경우 7.39m/s로 나타났다. 5% 비율이 가장 높은 평균 속도를 기록하며, 0% 비율 대비 64.23% 향상되었다. 이 결과는 자율주행 차량이 일정 비율 이상 도입될 경우 교통 흐름을 효율적으로 조절하며 전체적인 교통 혼잡을 개선하는 데 기여함을 보여주는 것으로 해석될 수 있다. 특히, 5% 비율에서 가장 높은 향상률을 보인 것은 자율주행 차량의 적절한 비율이 교통 흐름에 영향을 미칠 수 있음을 시사한다.

시뮬레이션을 통해 비 자율주행 차량만 존재하는 환경에서는 교차로 및 합류 구간에서 차량 정체 현상이 쉽게 발생했지만, 심층 강화학습을 적용한 자율주행 차량이 포함된 경우에는 학습할수록 차량 정체 현상이 비교적 적게 발생하는 것을 확인하였다.

III. 결론

본 논문에서는 도시 교통 문제 해결을 위한 가능한 대안 중 하나로 자율주행 차량의 도입을 제안하고 있다. 심층 강화학습 알고리즘인 PPO를 활용하여 자율주행 차량의 가감속을 조절하여, 전체 차량의 평균 속도를 Flow 시뮬레이션을 통해 확인하였고, 자율주행 차량이 교통 혼잡을 개선하는 데 도움이 되는 것을 보여주었다. 특히, 자율주행 차량의 비율이 5%일 때 가장 높은 평균 속도를 기록하였으나, 주어진 도로 환경과 시뮬레이션 파라미터에 영향을 받을 것으로 예상되어 추후 자율주행 차량의 비율에 따른 차량 흐름의 영향에 관해서 연구할 예정이다.

ACKNOWLEDGMENT

본 과제(결과물)는 2023년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다. (2021RIS-004)

참 고 문 헌

- [1] A. Zomarey, M. Rozhenko, "Impact of self-driving cars for urban development", *Foresight and STI Governance*, vol. 14, no. 1, p. 70-84, 2020.
- [2] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, "Proximal policy optimization algorithms", *arXiv preprint arXiv:1707.06347*, 2017.
- [3] M. Dian Khumara, L. Fauziyyah, P. Kristalina, "Estimation of urban traffic state using simulation of Urban mobility(sumo) to optimize intelligent transport system in Smart City", *2018 International Electronics Symposium on Engineering Technology and Applications (IES-ETA)*, pp. 163-169, 2018.
- [4] E. Liang, R. Liaw, R. Nishihara, P. Moritz, R. Fox, J. Gonzalez, K. Goldberg, I. Stoica, "Ray rllib: A composable and scalable reinforcement learning library", *arXiv preprint arXiv:1712.09381*, 2017.
- [5] M. Treiber, A. Kesting, "The intelligent driver model with stochasticity-new insights into traffic flow oscillations", *Transportation research procedia*, vol. 23, pp. 174-187, 2017.
- [6] E. Jacinto, F. Martinez, F. Martinez, "Navigation of Autonomous Vehicles using Reinforcement Learning with Generalized Advantage Estimation", *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 1, 2023.
- [7] M. Hardt, B. Recht, Y. Singer, "Train faster, generalize better: Stability of stochastic gradient descent", *In International conference on machine learning*, PMLR, pp. 1225-1234, 2016.