

차량 네트워크에서 상황에 따른 차별화 된 강화학습 모델의 필요성

유선종*, 김병찬, 김영부, 양태훈 오승민*

국립공주대학교

linebellu@smail.kongju.ac.kr*, smoh@kongju.ac.kr*

A Need for Reinforcement Learning Models according to Circumstances for Vehicular Networks

Seonjong Yoo, Byoungchan Kim, Youngboo Kim, Taehun Yang, Seungmin Oh

Kongju National University

요 약

차량 네트워크는 지속적으로 이동 중인 차량으로 데이터 통신을 제공하기 위한 기반 네트워크이다. 다양한 환경으로 이동 중이기 때문에, 차량은 상황에 따라 지속적 서비스를 보장받지 못한다. 이를 위해서 인공지능 기술이 접목되어 성능 향상을 기대하지만, 차량네트워크의 상황에 따른 환경변수들의 조합으로 처리량은 매우 상이 해 단일의 모델을 바탕으로 한 서비스는 최상의 결과를 보장받을 수 없다. 따라서 본 논문에서는 차량의 이동성으로 인해 발생하는 다양한 상황에 따른 통신성능을 강화학습을 통해 확인하고 환경변수에 따른 각기 다른 모델을 통해 학습을 진행해야 함을 보인다.

I. 서 론

차량 네트워크는 유/무선망을 통하여 차량 및 도로 등 인프라가 구축된 사물과 교통정보와 같은 정보를 교환하는 통신기술이다. 이 중 차량과 도로 인프라 간 통신(이하, V2I)은 도로에 위치한 차량 네트워크 노변장치(이하, RSU)와 차량의 무선통신이다. 이러한 네트워크를 통해 주행 간 수집되는 정보, 운전자가 필요로 하는 정보들을 원활하게 제공할 수 있어 차량 안전, 교통량 감소, 콘텐츠 소비 등에서 다양한 이점을 제공한다. 이러한 V2I 통신에서 차량은 매 순간 다른 속도와 이동방향, 경로 및 공존하고 있는 RSU들로 인해 토폴로지의 변동이 자주 발생한다[1].

현재 머신러닝을 활용한 네트워크 연구가 활발하게 진행되고 있다. 머신러닝에는 강화학습, 지도학습, 비지도학습이 있는데 [2]의 저자에 따르면 네트워크 연구에 지도학습이 60%, 강화학습이 40% 쓰이고 있다. 지도학습은 차량의 위치 예측, AP의 부하 예측과 개성이 활용되고, 강화학습은 시뮬레이션이 필요한 방대한 환경과, 자연보상, 확률적 의사 결정 처리로 여러 Wi-Fi 셀이 겹치는 구역과 다양한 네트워크가 공존하는 상황에 주로 쓰이고 있다[2].

무선랜 표준에서 상황에 따른 환경변수들의 조합과 차량, AP의 상태(Resource Unit)에 따라 처리량이 100%이상 달라지며 V2I 통신에서 Resource Unit은 시나리오에 따라 매우 다양하다[3].

본 논문에서는 V2I 통신에서 차량의 이동성으로 인해 발생하는 다양한 Resource Unit들에 따른 통신성능을 강화학습 중 Q-러닝을 통해 확인하고 상황에 따라 각기 다른 모델을 통해 학습을 진행해야 함을 보인다.

II. 본론

본 논문의 실험은 가로 1km, 세로 1km의 영역에서 차량이 구역을 통과하며 균등하게 배치되어있는 100개의 RSU와의 통신 감도와 RSU의 사용률에 따른 지연시간을 통하여 통신성능을 측정한다. 영역을 통과하는 차

량은 Q-러닝 기반으로 통신성능과 통과에 걸리는 시간 보상이 큰 방향으로 이동한다.

Q-러닝의 Q-table은 차량이 이동 간에 위치하는 각 좌표에 다음 행할 행동의 기대 보상 값을 저장하는 테이블로 수식(1)을 통해 학습과정에서 지속적으로 갱신된다. a 는 차량의 각 위치에서 할 수 있는 행동을 의미하며, 현 위치에서 앞으로 속도만큼 이동, 왼쪽으로 속도만큼 이동, 오른쪽으로 속도만큼 이동을 의미한다. $state$ 는 차량이 도착한 좌표를, R 은 $state$ 에서 a 를 수행할 때 기대되는 보상을 뜻한다. RD 는 강화학습에서 지나온 위치를 다시 지나올 때 보상을 줄이는 정도이며, LR 은 학습이 반복되면서 새로운 결과가 이전 결과를 대체하는 정도를 뜻한다.

$$Q-table[state, a] += LR \times (R + RD \times (R - previous\ Q-table[state - a])) \quad (1)$$

1. **For** 차량의 대 수
2. **While** 차량의 x좌표 < 1km
3. 수신감도 = 차량의 위치와 연결된 AP 사이의 거리
4. 지연시간 = 연결 된 AP의 기존 연결 차량 수
5. 경로점수 = 차량의 위치와 도착점까지의 거리
6. 보상 = 수신감도 ÷ 지연시간 + 경로점수
7. (a)차량의 현 위치에서 앞으로 속도만큼 이동 시 보상
8. (b)차량의 현 위치에서 왼쪽으로 속도만큼 이동 시 보상
9. (c)차량의 현 위치에서 오른쪽으로 속도만큼 이동 시 보상
10. (a~c)중 보상이 가장 큰 곳으로 이동

알고리즘 1. 강화학습 알고리즘

본 논문에서 기대 보상 값은 수식(2)와 같이 정의한다. ω 는 차량 1대가 수신 신호 세기가 최대를 유지하는 경로로만 이동했을 때의 통신성능으로 가장 이상적으로 기대되는 보상이며, α 는 수신 신호의 세기(차량과 RSU의 거리 제곱의 역수), β 는 연결된 차량과 RSU 간의 지연시간(RSU와 연결되어있는 전체 차량의 수)을 의미한다.

$$P = \frac{\alpha}{\beta} \times \frac{1}{\omega} \quad (2)$$

표 1 환경과 시나리오

환경	구분	값
	영역	1km * 1km
상황	RSU 수(개)	100
	RSU 위치	[0,0], [0,100], [0,200], ... [900,900]
	차량 속도(km/h)	[30, 40, 50, 60, 70, 80, 90, 100]
상황	차량 수(대)	[20, 40, 60, 80, 100]

시뮬레이션은 아래와 같은 조건에 따라 진행되며 동일한 환경에서 차량의 속도와 차량의 수를 변화시키며 각 시나리오 당 2000회 반복 학습하며 1km * 1km 영역의 좌표들의 기대 보상 값을 수식(1)을 통해 갱신한다. 그림 1, 2는 시뮬레이션 결과 최종 생성된 Q-table을 따라 이동하는 차량의 수와 속도에 따른 통신성능이다.

- 차량은 랜덤한 가로축 0km 지점에서 출발하여 가로축 1km 지점에 도착하면 종료된다.
- 차량은 동일한 위치에 동시간에 있을 수 없다.
- 차량은 앞으로 속도만큼 이동, 왼쪽으로 속도만큼 이동, 오른쪽으로 속도만큼 이동 이 3가지의 움직임을 가질 수 있다.
- 수신 신호의 세기는 (이를 결정하는 가장 영향력 있는 요소인) 차량과 RSU의 거리로 대체한다.
- 차량은 수신 신호 세기가 가장 좋은 RSU와 연결된다.
- 연결된 차량과 RSU 간의 지연시간은 (이를 결정하는 가장 영향력 있는 요소인) RSU와 연결되어있는 전체 차량의 수로 대체한다.

ω (이상적으로 기대되는 보상)은 속도와 상관없이 오차범위 5% 이내의 비교적 일정한 값인 반면 그림 1에 따르면 차량의 수와 속도에 따라 전혀 다른 성능을 보여준다. 특히 적은 수의 차량인 상황과 많은 수의 차량인 상황에서 확연히 다른 성능 경향을 보인다.

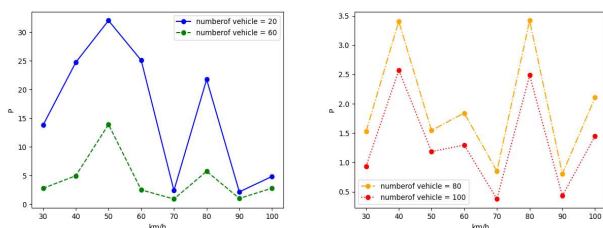


그림1 저밀도 차량(좌)과 고밀도 차량(우) 환경에서 속도별 성능 경향

그림 2에 따르면 차량의 수가 늘어날수록 성능이 대체로 저하되는 것을 확인할 수 있다. 60km/h 이하의 속도에서는 차량 수 60대를 넘어가면서 급격하게 성능이 저하된 반면, 70km/h 이상부터는 차량 수 40대~80대까지 성능이 비교적 일정하다.

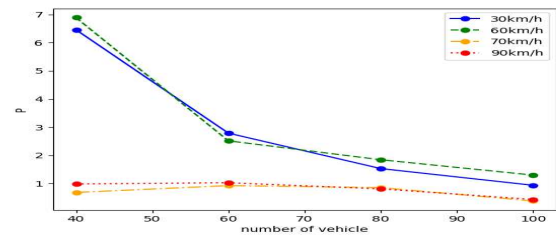


그림2. 저속과 고속에서의 성능 경향

실험결과 ω 는 대체로 일정한 반면, 실험의 상황(속도와 차량 수)에 따라서 다른 성능과 특성을 보인다.

III. 결론

본 논문에서는 차량 네트워크 중 V2I에서 차량의 통신성능을 보상으로 한 강화학습을 진행하였다. V2I의 다양한 환경 중 차량의 속도와 차량의 수에 따른 통신성능 결과는 상황에 따라 상이했으며 이에 따라 차량 네트워크의 각 시나리오를 구분하여 각기 다른 모델을 통해 학습을 진행해야 한다.

실험의 통신성능은 가장 영향이 큰 요인들로 대체하여 개략적으로 도출하였다. 실제 차량 네트워크 환경에서는 보다 많은 요인들이 통신성능에 영향을 끼칠 것으로 향후 연구과제에서 통신성능에 실제 모델을 도입이 필요하다. 또한 성능 경향의 원인 분석과 RSU의 배치를 포함한 매개변수들을 다양화하여 각기 다른 상황에 강화학습 모델의 분리 기준을 정의하고자 한다.

ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원 받아 수행된 연구임(No. NRF-2020R1C1C1010692)

참 고 문 헌

- [1] K. Mershad and H. Artail, "SCORE: Data Scheduling at roadside units in vehicle ad hoc networks," proc 2012 19th International Conference on Telecommunications (ICT), Jounieh, Lebanon, pp. 1-6
- [2] J. Kaur, M. A. Khan, M. Iftikhar, M. Imran and Q. Emad Ul Haq, "Machine Learning Techniques for 5G and Beyond," IEEE Access, vol. 9, pp. 23472-23488, 2021
- [3] S. Szott et al., "Wi-Fi Meets ML: A Survey on Improving IEEE 802.11 Performance With Machine Learning," IEEE Communications Surveys & Tutorials, vol. 24, no. 3, pp. 1843-1893, thirdquarter 2022