

심층강화학습을 이용한 Cellular-V2X 요구사항 적응형 혼잡제어 알고리즘

양우열, 조한신

한밭대학교 전자공학과

30211173@edu.hanbat.ac.kr, hsjo@hanbat.ac.kr

Cellular-V2X Requirements Adaptive Congestion Control Algorithm Using Deep Reinforcement Learning

Wooyeol Yang, Han-Shin Jo,

Hanbat National University

요약

Cellular-Vehicle to everything (C-V2X)는 Long Term Evolution 기반의 V2X통신을 지칭하며, 해당 통신 기술의 요구 사항으로는 메시지 수신 신뢰도와 통신 범위 등이 있다. 그러나 C-V2X는 고밀도 상황에서 높은 혼잡도에 의해 해당 요구 사항을 만족시키지 못하며, 이를 위해 V2X관련 표준 기관은 혼잡제어 알고리즘을 발표했다. 그러나 이와 같은 기존 혼잡제어 알고리즘은 C-V2X 요구 사항을 고려하지 않은 채 설계되었기 때문에 요구 사항을 만족시키기 어렵다. 따라서, 본 논문에서는 C-V2X 요구 사항을 만족시키기 위해 심층강화학습을 이용한 중앙 집중식 혼잡제어 알고리즘을 제안한다. 제안하는 알고리즘에서 에이전트는 기지국으로 네트워크의 혼잡도를 측정하여 모든 차량의 전송 주기를 일괄적으로 제어한다. 본 논문에서 이용한 심층강화학습 알고리즘은 Deep Q-Networks (DQN)와 Double DQN이며, 시뮬레이션 결과 두 알고리즘 모두 목표로 하는 C-V2X 요구 사항을 만족시킨다.

I. Cellular-Vehicle to everything (C-V2X)

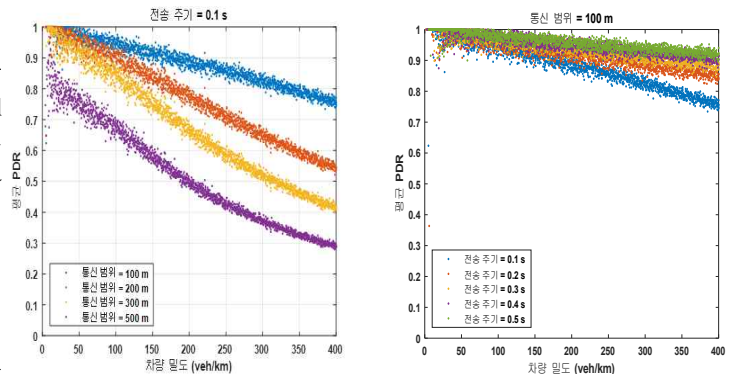
C-V2X는 Long Term Evolution 기반의 V2X통신을 지칭하며, 여기서 V2X란 차량이 유·무선망을 통해 다른 차량 및 도로 등 인프라가 구축된 사물과 정보를 교환하는 것 또는 그 기술을 말한다. 국제 표준화 단체 3rd Generation Partnership Project (3GPP)에서는 이러한 C-V2X 요구 사항을 기술 문서에서 제안하고 있으며, 그 예로 메시지 수신 신뢰도(Packet Delivery Ratio, PDR)는 통신 시나리오에 따라 80~95 %, 협력 인식 메시지의 통신 범위는 300~500 m로 제안하고 있다[1].

그러나 C-V2X는 고밀도 상황에서 통신 성능이 급격히 저하되기 때문에 PDR, 통신 범위와 같은 요구 사항을 만족시키기 어렵다. 이러한 문제점을 해결하기 각 V2X와 관련된 표준 기관들은 전송 전력제어, 전송 주기 제어와 같은 혼잡제어 알고리즘을 발표했지만, 이와 같은 기존 혼잡제어 알고리즘은 C-V2X의 요구 사항을 고려하지 않은 채 설계되었기 때문에 요구 사항을 효율적으로 만족시키지 못한다. 따라서 본 논문에서는 심층강화학습을 이용한 C-V2X 요구 사항 적응형 중앙 집중식 혼잡제어 알고리즘을 제안한다. 제안하는 알고리즘의 혼잡제어 기술은 전송 주기 제어를 이용하여 차량 간 패킷 충돌 현상을 조절한다. 이를 통해 목표로 하는 통신 요구사항을 만족시킬 수 있으며, 설정한 목표 통신성능에 따라 상승 시키거나 저하시킬 수 있다.

II. 전송 주기 및 통신 범위설정에 따른 차량 네트워크 성능 변화

본 논문에서 활용한 C-V2X의 요구 사항은 PDR과 통신 범위이다. 다음 그림 1은 전송 주기, 통신 범위에 따른 차량 네트워크의 평균 PDR을 보여준다. 그림 1. (a)는 네트워크 내 차량들의 전송 주기를 0.1 s로 고정시킨 결과이다. 결과에서 확인할 수 있듯이, 각 차량이 커버해야 할 통신 범위가 넓어질수록 평균 PDR이 낮아지는 것을 알 수 있다. 이러한 결과는 차량 간 송 수신 거리가 멀어짐에 따라 경로 손실에 의해 신호대 잡음 간섭비가 낮아지기 때문이다.

다음 그림 1. (b)는 통신 범위를 100 m로 설정하여, 차량 간 전송 주기 설정



(a) 통신 범위에 따른 평균 PDR

(b) 전송 주기에 따른 평균 PDR

그림 1. 변수에 따른 차량 네트워크의 평균 PDR 변화

에 따른 네트워크의 평균 PDR을 보여준다. 결과에서와 같이 전송 주기를 늘릴수록 패킷 충돌 현상이 줄어들기 때문에, 평균 PDR결과가 높아지는 것을 확인할 수 있다. 그러나 네트워크가 혼잡하지 않은 상황에서 전송 주기를 늘리는 것은 메시지 지연을 야기하기 때문에, C-V2X 메시지 수신 신뢰도를 만족시킬 수 있는 적절한 전송 주기 선택이 필요하다.

최종적으로 목표로 하는 통신 범위 내 요구되는 평균 PDR 성능을 만족시키기 위해서는 통신 네트워크 상황에 따른 전송 주기 제어가 필요하다. 이는 네트워크 혼잡도에 따른 전송 주기의 조합 최적화 문제로 전환되며 높은 복잡성을 갖는다. 따라서, 본 논문에서는 이러한 조합 최적화 문제를 해결하기 위해 심층강화학습 알고리즘 중 Deep Q-Networks (DQN)와 Double DQN (DDQN)을 이용한다[2].

III. 심층 강화학습을 이용한 C-V2X 요구 사항 적응형 중앙 집중식 혼잡제어 알고리즘

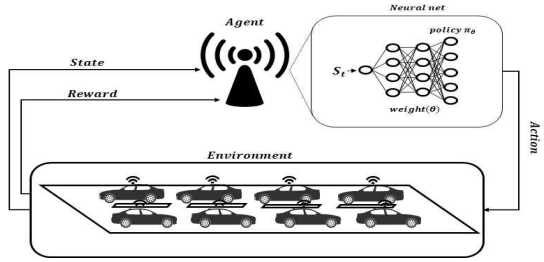


그림 2. 알고리즘의 강화학습 모델

제안하는 알고리즘에서 기지국은 심층강화학습을 통해 훈련되며, 각 차량은 도로의 혼잡도를 나타낼 수 있는 100 m 이내 노드 수를 기지국에게 전달한다. 이후 전달받은 정보를 토대로 기지국은 네트워크 내 모든 차량의 전송 주기를 변화 시키며, 이때 전송 주기는 모든 차량이 동일하게 부여받는다. 전송 주기는 기지국에 설정된 C-V2X 요구 사항 목표에 맞추도록 변화되며 이는 그림 2와 같은 강화학습 모델로 정의된다.

본 연구에서 적용된 심층강화학습 알고리즘인 DQN은 타겟 네트워크의 가중치 (θ')와 코럴 네트워크의 가중치 (θ)를 최적화하여, 에이전트의 정책 (π_θ)을 최적화시키는 알고리즘이다. 여기서 π_θ 는 입력으로 t 시점의 관측치 (State, S_t)를 입력으로 받아 행동 (Action, A_t)의 확률을 출력하는 함수이다. 최적화된 π_θ 은 보상 (Reward, R_t)을 최대화한다. 제안하는 알고리즘에서 S_t 는 100 m 이내 평균 노드 수로 설정했으며, A_t 는 전송 주기로 0.1~0.5 s의 5개의 단계로 설정했다. 최종적으로 R_t 는 C-V2X 요구 사항을 만족시키기 위해 다음과 같이 설정했다.

$$R_t = -|PDR_{Target} - OUTPUT| \quad (1)$$

여기서 PDR_{Target} 은 목표로 설정한 네트워크의 통신 범위 내 평균 PDR이며, $OUTPUT$ 은 실제 측정된 평균 PDR값이다.

DQN알고리즘은 다음 수식과 같이 신경망의 손실 함수를 Mean Squared Error (MSE)로 정의한다.

$$L = \frac{1}{M} \sum_{i=1}^M (y_i + Q(S_i, A_i | \theta))^2 \quad (2)$$

위의 M 은 experience replay memory에서 추출한 mini-batch 사이즈이며, $Q(S_i, A_i | \theta)$ 는 State에서의 Action의 가치를 나타내는 Q-value로 코럴 네트워크의 값이다. y_i 는 목표 값으로 에피소드 종료 유/무에 따라 다음 식으로 정의된다.

$$\begin{cases} R_i & \text{for } S_i = S_T \\ R_i + \gamma \max_{A'} Q'(S_i, A' | \theta') & \text{for } S_i \neq S_T \end{cases} \quad (3)$$

식에서 γ 는 감가율로 0과 1사이 값이며 T 는 에피소드 종료 시점을 의미한다. $Q'(S_i, A' | \theta')$ 는 타겟 네트워크에서 추출한 Q-value이다. DQN과 달리 DDQN은 다음의 수식으로 y_i 가 정의된다.

$$\begin{aligned} A_{\max} &= \arg \max_{A'} Q(S_i, A' | \theta) \\ y_i &= R_i + \gamma Q'(S_i, A_{\max} | \theta') \end{aligned} \quad (4)$$

위와 같이 목표 값의 업데이트 수식을 다르게 하여 DQN과 달리 DDQN은 학습의 편향을 줄여주는 효과를 낸다.

IV. 시뮬레이션 결과

표 1. 시뮬레이션 파라미터

파라미터	값
메시지 크기	300 byte
도로 길이	1 km
대역폭	10 MHz
안테나 이득 (G_t, G_r)	3 dB
목표 PDR (PDR_{Target})	0.95, 0.85, 0.7
통신 범위	100, 200 m
MCS index	3

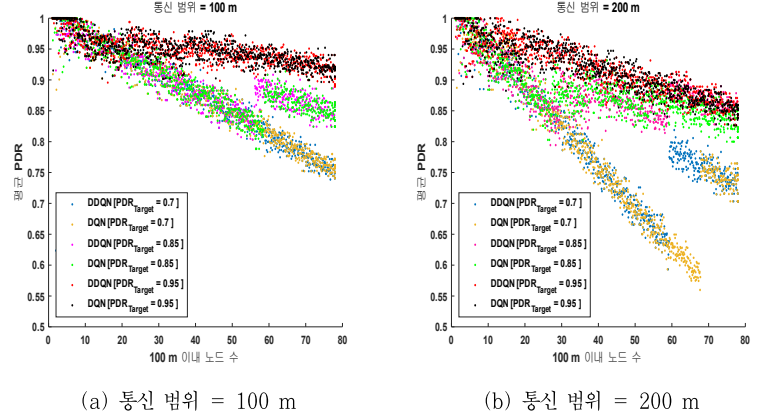


그림 2. 통신 범위와 목표 PDR에 따른 학습결과

시뮬레이션은 표 1의 파라미터를 이용했으며, DQN과 DDQN으로 각각 학습한 기지국의 성능을 평가한다. 시뮬레이션 평가는 차량 네트워크의 평균 PDR을 이용한다. 다음 그림 2는 통신 범위와 PDR_{Target} 에 따른 두 가지 심층강화학습 알고리즘의 학습 결과를 보여준다. 그림 2. (a)는 통신 범위가 100 m로 설정했기 때문에 (b)보다 전반적으로 높은 수준의 평균 PDR을 보인다. 두 결과에서 모두 PDR_{Target} 이 높을수록 기지국은 전송 주기를 0.5 s에 가깝게 이용했으며, PDR_{Target} 낮을수록 0.1s 가까운 전송 주기를 이용하였다. 또한 통신 범위가 100 m에서 200 m로 증가함에 따라 기지국은 같은 관측치에서 더욱 긴 전송 주기를 이용했다. 시뮬레이션 결과에서 DQN과 DDQN의 성능 차이가 크지 않은 이유는 시뮬레이션 환경에서 관측치에 대한 에이전트의 행동 양상이 뚜렷하기 때문으로 분석된다.

V. 결론

본 논문에서는 심층강화학습을 이용하여 C-V2X 요구 사항에 적응적인 혼잡 제어 알고리즘을 제안한다. 제안하는 알고리즘에서 에이전트는 기지국으로 네트워크 내 모든 차량의 전송 주기를 제어한다. 시뮬레이션 결과, 기지국은 네트워크 혼잡도에 따라 통신 범위 내 PDR_{Target} 을 만족시키기 위해 유연한 전송 주기 제어를 보였으며, 이를 통해 PDR_{Target} 을 만족시키는 것을 확인할 수 있었다. 추후 심층강화학습을 이용한 분산 혼잡 제어 알고리즘을 연구할 계획이다.

참 고 문 헌

- [1] Study on LTE support for Vehicle to Everything (V2X) services, 3GPP TR 22.885 V14.0.0, Dec. 2015.
- [2] R. S. Sutton and A. G. Barto, Reinforcement learning: An introduction. MIT press, 2018