

NR V2X에서의 DQN 기반 자원 할당 Mode 2 최적화

정현수, 신은지, 김태형*
국민대학교

jhs6339@kookmin.ac.kr, tldsdmswl7287@kookmin.ac.kr, *th.kim@kookmin.ac.kr

Optimization for Resource Allocation Mode 2 based on DQN for NR V2X

Hyeonsu Jeong, Eunji Shin, Taehyoung Kim*
Kookmin University

요약

본 논문은 V2X Mode2 통신 기술에서 다양한 요구 조건을 만족하기 위해 DQN(Deep Q-Network)를 활용한 자원 할당 최적화 방법을 제시한다. DQN 알고리즘을 통해 송신 차량과 수신 차량 간의 효율적인 자원 배분을 실시간으로 최적화하고, 차량 간의 상호작용을 고려하여 네트워크 전체 데이터 전송률 합을 최대화하였다. 이를 통해 빠르게 변동하는 차량 간 통신 환경에서도 안정적인 자원 할당이 가능하고, 고주파 대역에서 발생할 수 있는 통신 불안정성을 해결하며 효율적인 자원 분배를 실현할 수 있음을 실험적으로 입증하였다. 향후 다양한 통신 환경에서도 활용 가능성을 보여준다.

I. 서론

최근 자율주행 및 스마트 시티 기술의 급격한 발전에 따라, 차량 간 실시간 정보 교환이 필수적인 요소로 자리 잡고 있다. V2X 통신은 다양한 주체 간 실시간 정보 교환을 가능하게 하고 자율주행 차량이 눈앞에 보이지 않는 위험을 미리 감지하고 대처할 수 있게 해주는 핵심 기술로 교통 시스템의 안전성과 효율성을 향상시키고 있다. 그러나, 고속으로 이동하는 차량 간 통신 환경에서는 높은 변동성과 제한된 주파수 자원으로 인해 통신 품질이 불안정해지는 문제가 발생할 수 있다. 이에 따라 V2X 통신 시스템에서는 실시간으로 자원을 효율적으로 할당하고, 고주파 대역에서도 안정적인 통신을 보장할 수 있는 기법이 필요하다.

본 논문에서는 V2X Mode2 통신의 효율성을 향상시키기 위해 심층강화학습(DQN) 기법을 사용한다. DQN 기법은 기존의 강화학습 방법을 심층 신경망으로 확장한 기법으로, 빠르게 변동하는 환경에서도 최적의 자원 할당을 학습할 수 있다. 따라서, DQN 기법을 활용하여 차량 간 자원 할당을 최적화하는 방법을 제시한다.

II. NR V2X RA Mode 2

NR V2X RA Mode 2는 기지국 개입 없이 UE 간 자원 할당을 수행하는 방식이다.[1] UE는 송신 전 미리 정의된 센싱 윈도우동안 모든 자원에 대해 RSRP 및 간섭 레벨을 측정하여 현재 사용중인 자원을 탐지한다. 측정된 RSRP가 사전 설정된 임계값을 초과하면 자원을 후보 풀에서 제외하고 후보 자원 중 하나를 선택하여 할당한다. 본 연구에서는 자원 수를 N=4로 고정하고 각 자원에 대해 주변 UE로부터 수신되는 간섭 전력을 -90~ -80dBm 범위에서 랜덤으로 생성하였다. 임계값 Threshold=-87dBm으로 적용하여 UE가 각 자원의 간섭 전력을

컬로 측정하여 설정된 임계값 미만인 자원만 전송 후보로 선별하도록 RA Mode 2를 약식 구현하였다. 이렇게 선별된 후보 자원 중 하나를 선택해 실제 전송을 수행하여 자원 할당을 실현하였다. 이와 같이 센터 제어 신호 없이 Mode 2의 알고리즘을 구현하였다.

III. 제안하는 DQN 기반 RA Mode2[2]

본 논문에서는 V2X 통신 환경에서 DQN 기반 자원 할당 기법을 적용하기 위한 네트워크 모델을 제시한다. 심층강화학습(DQN)은 강화학습과 최적의 심층신경망을 결합한 기법으로, 핵심 목표는 보상의 누적 합인 Q 값을 최적화하는 것이다. 이를 통해 에이전트는 최적의 행동을 선택하게 된다. 최적의 Q 값을 찾기 위해서는 Bellman equation에 기반하여 다음과 같이 Q 함수를 정의할 수 있다.[3]

$$Q^*(s, a) = \mathbb{E}_s [r + \gamma \max_a Q^*(s', a') | s, a]$$

s는 상태, a는 행동, r은 보상을 나타내며, γ 는 감쇄인자이다.

1) 상태(state) s: 송신단과 수신단 사이의 채널 이득 $g(\theta_{m,k,l})$, 각도 $\alpha_{m,k,l}$, 데이터 요구량 R_{req}

2) 행동(action) a: 자원 할당 행렬 A

3) 보상(reward) r: 전체 네트워크 데이터 전송률의 합을 최대화하도록 한다. 수신 차량의 데이터 전송률 요구 조건을 만족하지 못하면 차량마다 -1의 페널티를 준다. $I(x)$ 는 x가 참일 때 1, 아니면 0인 함수이다.

$$\sum_{k=1}^K [R_k(A) - I(R_k(A) < R_{req})]$$

송신차량 M 대와 수신 차량 K 대가 존재하는 네트워크 상황을 가정하여 채널 모델을 구성한다.

표 1. 시뮬레이션 파라미터

Parameter	Value
Bandwidth	4 [MHz]
thermal noise density	174 [dBm]
Pathloss coefficient	3.76

$$h_{m,k} = \sum_{k=1}^L \alpha_{m,k,l} g(\theta_{m,k,l}) d^{-3.76}$$

L 은 경로의 수, d 는 차량이 움직임에 따라 변하는 거리, -3.76 을 path loss 로 설정해 거리에 따른 path loss 를 적용하였다. 배열응답 벡터는 다음과 같이 나타난다.

$$g(\theta_{m,k,l}) = [1, e^{j\pi \sin \theta_{m,k,l}}, \dots, e^{j(N-1)\pi \sin \theta_{m,k,l}}]^T$$

m 번째 송신 차량의 자원 할당 상태와 네트워크 전체 자원 할당 상태는 다음과 같이 정의한다.

$$\begin{aligned} A &= [A_{1,1,1}, A_{1,1,2}, \dots, A_{1,1,A}, A_{1,2,1}, A_{1,2,2}, \dots, A_{m,k,A}] \\ A &= [A_1, \dots, A_M] \end{aligned}$$

이때, 벡터의 각 요소는 m 번째 송신 차량의 k 번째 수신 차량에 A 번째 주파수 할당 여부에 대한 변수로 정의된다.

A_1 은 단말#1 이 Mode 2 동작에 따라 RSS(resource sensing and selection)를 통해 판단한 가용 자원을 의미한다. 이에 따라 수신 차량에 따른 데이터 전송률은 다음과 같이 나타낸다. 송신 차량 빔포밍 벡터를 포함한다.

$$R_k(A) = \log_2 \left(1 + \frac{\sum_{m=1}^M \sum_{s=1}^S P_{m,k,s} |\hat{h}_{m,k}^H f_{m,k}|^2}{\sum_{m=1}^M \sum_{s=1}^S \sum_{j \neq k}^K P_{m,j,s} |\hat{h}_{m,k}^H f_{m,j}|^2 + \sigma^2} \right)$$

$$f_{m,k} = \frac{\hat{h}_{m,k}}{|\hat{h}_{m,k}|}$$

DQN 에이전트가 상태 s 를 입력으로 받아 가능한 자원 할당 행렬 A 중 최적의 할당을 선택하도록 학습한다. 에이전트는 ϵ -탐욕 정책을 통해 행동을 선택하고 그 결과 얻은 전송률 합을 최대화하는 방향으로 학습을 함으로써 전체 네트워크 데이터 전송률을 극대화하는 자원 할당 정책을 획득하게 된다.

IV. 실험 결과

제안된 DQN 알고리즘을 이용하여 V2X 통신에서의 자원 할당 효율성을 검증하기 위해 6차선 고속도로 상황에서 M(송신차량) 5 대, K(수신차량) 5 대의 구성하고 심층 강화 학습 네트워크는 은닉층 4 개로 구성되며, 각 은닉층 당 뉴런수가 300이고 입실론(ϵ)은 0.1로 설정하였다.

그림 1에서 볼 수 있듯, 제안한 DQN 기반 자원 할당은 무작위 할당에 전송전력 평균 약 4.5 배 전송률 향상을 달성하였다. 이론적 최적 알고리즘과 비교했을 때 약 65~70% 수준의 성능을 유지하여 최소한의 성능 손실만으로 효율적인 자원 활용이 가능함을 확인하였다. 이러한 결과는 DQN 에이전트가 수신 차량들의 데이터 요구 조건의 상황을 고려하여 데이터 전송률의 최대 합을 얻는 것을 알 수 있었다.

그림 2에서 x 축은 송신기 활성화 개수를 1 대부터 5 대로 설정하고, y 축은 스케줄링 수로 설정하였다. 송신기 수에 따른 스케줄링 횟수 편차가 약 10% 이내로 비교적 균일하게 분포함을 확인할 수 있었다. 네 개의 자원만을 이용하는 환경에서도 스케줄링은 각 수신 차량의 개별 데이터 요구량을 충족시키면서 전체 전송률을 극대화하도록 자원을 효율적으로 배분하는 것을 확인했다. 이는

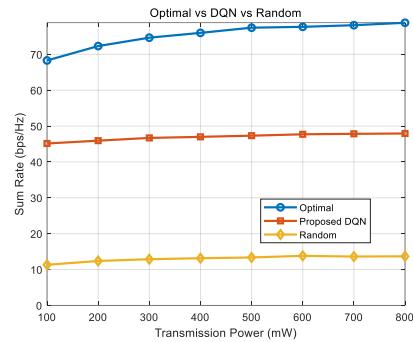


그림 1. DQN 적용과 최적 성능 비교 그래프

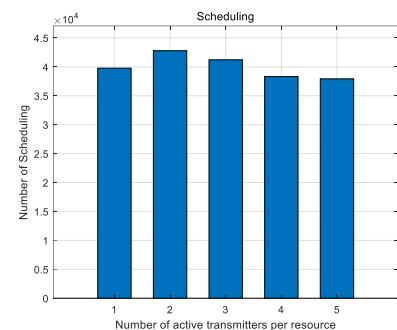


그림 2. 송신기 활성화 개수에 따른 스케줄링 분포

DQN 에이전트가 부하 조건에서도 안정적으로 자원을 탐색, 할당하며 정책 대비 일관된 스케줄링 기회를 보장함을 의미한다.

V. 결론

본 논문에서는 V2X 통신에서 차량 간 실시간 정보 교환의 효율성을 높이기 위해 DQN 기법을 활용하여 자원 할당 최적화하는 방법을 제시하였다. 고주파 대역 통신에서도 효율적인 자원 할당이 가능함을 실험적으로 입증하였다. 이러한 기법은 고속 이동 환경에서도 실시간으로 자원 할당을 최적화할 수 있어, 자율주행 차량의 통신 효율성 향상에 기여할 것으로 기대된다.

ACKNOWLEDGMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No. RS-2024-00397789, Development and Standardization of 6G Radio Transmission Standard Technology)

참고 문헌

- [1] T. Kim, Y. Kim, M. Jung and H. Son, "Intelligent Partial-Sensing-Based Autonomous Resource Allocation for NR V2X," in IEEE Internet of Things Journal, vol. 11, no. 2, pp. 3144–3160, 15 Jan. 15, 2024.
- [2] 3GPP TS 38.214, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Physical layer procedures for data (Release 18)," V18.6.0 , March, 2025
- [3] H. Eun-gyu and C. Chang-bok, "Reinforcement Learning Model Implementation for Trading Prediction Using DQN," Journal of KIIT, vol. 17, no. 4, pp. 1–8, 2019.