

DDQN을 활용한 NR-V2X 모드2의 동적 리소스 예약 조절 기법

김성현*, 김채현*, 양영준*, 김동균*

*경북대학교

*sunghyunkim@knu.ac.kr, *hyeon917@knu.ac.kr, *youngj719@knu.ac.kr, *dongkyun@knu.ac.kr

A DDQN-based Control Method for Dynamic Resource Reservation in NR-V2X Mode 2

*Kim Sung Hyun, *Kim Chae Hyeon, *Yang Young Joon, *Kim Dong Kyun

*Kyungpook National University

요약

본 논문은 3GPP Release 16 표준에 정의된 NR-V2X의 차량 대 차량(V2V) 통신 환경을 다룬다. V2V 통신 모드 2에서는 반-영구 스케줄링(SPS) 방식을 통해 자원을 예약하며, 이때 RC(Reserved Counter)와 SRRI(Sidelink Reserved Resource Interval) 파라미터가 사용된다. 하지만, 차량의 주행 환경 및 트래픽 패턴 변화로 인해 데이터 크기가 동적으로 변할 경우, 사전에 고정된 RC 값으로 인해 리소스 낭비 및 활용률 저하 문제가 발생한다. 본 논문에서는 이러한 문제를 해결하기 위해, 강화학습 기법 중 하나인 DDQN(Double Deep Q-Network)을 활용하여 주변 환경 변화에 따라 동적으로 리소스 예약을 조절하는 이론적 접근법을 제안한다.

I. 서론

5G 네트워크 기반의 NR-V2X 기술 표준화를 위해 3GPP는 Release 16 문서를 통해 V2V 통신 방식을 정의 하였다[1]. V2V 데이터 통신을 위해 모드2의 리소스(Sidelink) 예약 기법을 Release 16에서 소개하였다. 이 중 모드 2는 기지국(eNodeB)의 개입을 최소화하며 차량이 자율적으로 통신 자원을 선택하는 방식으로, SPS(Semi-Persistent Scheduling) 기법을 사용한다. 차량은 데이터 전송이 필요할 때 RC(Reserved Counter)와 SRRI(Sidelink Reserved Resource Interval)값을 기반으로 미래의 특정 시점까지 자원을 미리 예약한다[2].

하지만, 이러한 정적 자원 예약 방식은 실시간 트래픽 수요 변화에 유연하게 대응하기 어렵다는 한계가 있다. 예를 들어, 데이터 전송이 조기에 완료되거나 전송할 데이터의 크기가 급격히 줄어들었음에도, 초기에 설정된 RC 값만큼 자원이 불필요하게 예약되어 다른 차량이 사용하지 못하는 비효율이 발생한다. 이러한 문제는 전체 네트워크의 리소스 활용률을 저하시키는 주된 원인이 된다.

본 논문에서는 이러한 문제를 해결하기 위해 강화학습을 적용하고자 한다. 특히, DDQN(Double Deep Q-Network) 에이전트를 통해 주변 통신 환경을 인지하고, 최적의 RC 값을 실시간으로 결정하여 자원 낭비를 최소화하고 활용률을 극대화하는 기법을 제안하며, 2장에서는 제안기법을 서술하고 3장 결론으로 마무리 한다.

II. 본론

2.1 문제정의

NR-V2X 모드 2에서 차량은 패킷이 발생하면 eNodeB로부터 할당받은 자원 풀 내에서 특정 자원을 선택하고, RC 값을 설정하여 SRRI 간격으로 해당 자원을 예약한다. 이 SPS 방식은 데이터 충돌과 혼잡을 줄이는 데 효과적이다.

하지만 차량 밀도나 주변 환경 변화로 인해 통신 트래픽의 특성은 계속해서 변한다. 예를 들어, 안전 메시지의 생성 주기가 바뀌거나 전송 데이터의 크기가 감소하면, 기존에 예약된 자원이 과도하게 되어 다른 차량의

통신 기회를 박탈하고 채널 낭비를 유발한다. 본 논문에서는 '정적 RC 값'과 '동적 트래픽 수요' 사이의 불일치 문제를 해결하는 것을 목표로 한다.

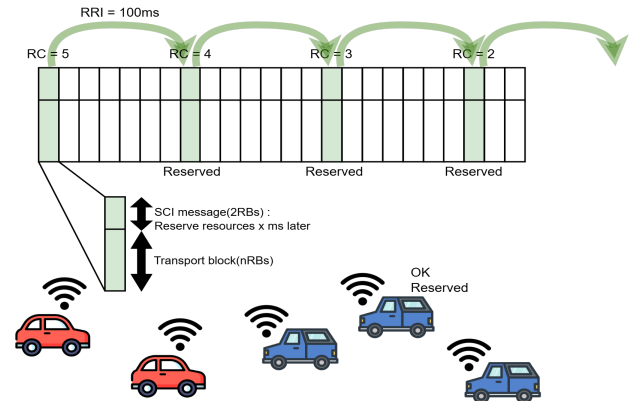


그림 1 NR-V2X 모드2 RC와 SRRI System

2.2 제안 기법: DDQN 기반 동적 RC 조절

본 논문에서는 DDQN을 이용하여 RC 값을 최적화하는 기법을 제안한다. DDQN은 기존 DQN(Deep Q-Network)이 갖는 Q-가치 과대평가(Overestimation) 문제를 완화하여 학습 안정성을 높인 알고리즘이다[3]. 제안하는 강화학습 모델의 주요 파라미터는 다음과 같이 정의한다.

● State (상태): 에이전트(차량)가 최적의 행동을 결정하기 위해 관찰하는 정보의 집합. 현재 통신 환경을 정확히 나타낼 수 있는 파라미터들로 구성된다.

· $S = \{RSSI \text{ (Received Signal Strength Indicator), 현재 PDR (Packet Delivery Ratio), 현재 설정된 RC 값, 잔여 데이터 버퍼 크기}\}$

● Action (행동): 에이전트가 특정 상태(State)에서 취할 수 있는 정책적 결정. 본 문제에서는 RC 값을 동적으로 조절하는 것이 목표이다.

· $A = \{RC \text{ 값 1 감소, RC 값 유지, RC 값 1 증가}\}$ 와 같이 정의하거나,

특정 범위 내의 RC 값 중 하나를 직접 선택하도록 설계할 수 있다.

● Reward (보상): 에이전트가 취한 행동에 대한 성능을 평가하는 지표. 보상 함수는 리소스 활용률을 높이고 패킷 전송 성공률을 보장하는 방향으로 설계되어야 한다.

$$R = (w1 * PDR) - (w2 * Wasted_Resources)$$

여기서 $w1$ 과 $w2$ 는 가중치이며, PDR은 높을수록 긍정적 보상을, 예약되었지만 사용되지 않은 자원의 양을 의미하는 Wasted_Resources는 클수록 부정적 보상(패널티)을 제공한다.

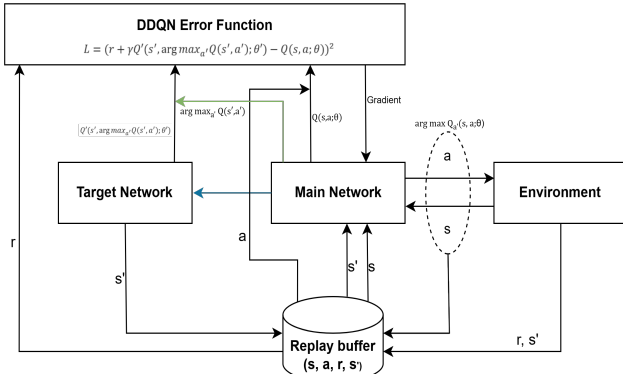


그림 2 DDQN

DDQN 에이전트는 '상태'를 관찰하고 '행동'을 수행한 뒤, 그 결과로 '보상'을 받으며 학습을 진행한다. 이 과정을 반복하여, Main Network와 Target Network 간의 오차(Loss)를 최소화하는 최적의 정책(Policy), 즉 각 상황에 맞는 최적의 RC 값을 찾는 규칙을 학습하게 된다. 이를 통해 데이터 크기 변화에 유연하게 대응하고 리소스 활용률을 극대화할 수 있다.

2.3 실험 및 결과

본 제안 기법의 실험은 NS-3를 이용하여 PDR(Packet Delivery Ratio)와 리소스 활용률에 대해 기존 고정 RC와 비교 분석 하였으며 실험 환경은 아래의 표1과 같다.

파라미터(Parameter)	값(Value)
차량 수	10 ~ 100(10대씩 증가)
시뮬레이션 시간	300s
주파수 대역	5.9GHz
대역폭	10MHz
SRRI	100ms
전송 전력	23dBm
패킷 크기	200 ~ 500 Bytes

표 1 NS-3 시뮬레이션 파라미터

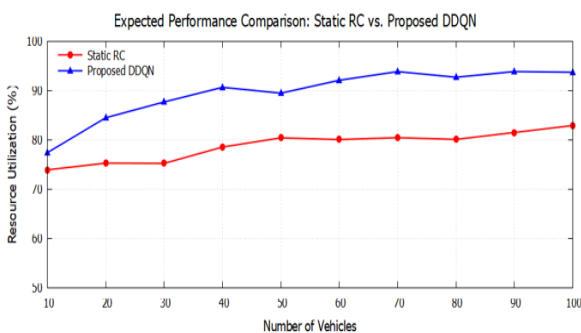


그림 3 리소스 활용률

그림3은 제안 기법이 차량 수가 증가해도, 제안하는 DDQN 기법은 85% 이상의 높은 리소스 활용률을 꾸준히 유지한다. 반면, 고정된 RC를 사용하는 기존 방식은 약 80% 수준에서 정체된다. 이는 DDQN 에이전트가 변화하는 트래픽 수요에 맞춰 불필요한 자원 예약을 효과적으로 줄여 자원 낭비를 최소화했음을 의미한다.

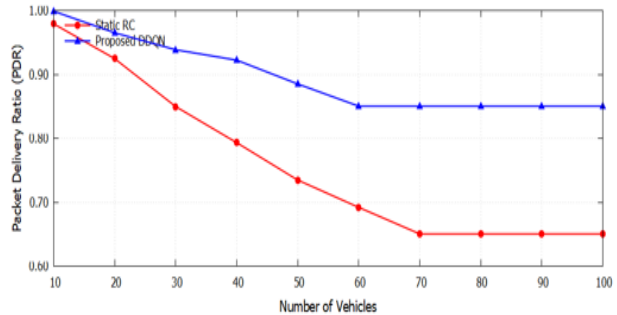


그림 4 PDR(Packet Delivery ratio)

그림4는 차량 수가 늘어날수록 두 기법 간의 PDR 성능 차이는 더욱 두드러진다. 기존 방식의 PDR은 65%까지 급격히 하락하는 반면, DDQN 기법은 85% 수준을 안정적으로 방어한다. 이는 DDQN이 유연한 자원 관리를 통해 패킷 충돌을 줄이고 통신 안정성을 크게 향상시켰음을 보여준다.

III. 결론

본 논문에서는 NR-V2X V2V 통신의 모드 2 환경에서 고정된 RC와 SRRI 값으로 인해 발생하는 리소스 낭비 및 활용률 저하 문제를 지적했다. 이에 대한 해결책으로, 강화학습 기반의 DDQN 알고리즘을 적용하여 주변 환경과 트래픽 수요에 맞춰 RC 값을 동적으로 조절하는 기법을 제안했다. 상태, 행동, 보상을 명확히 정의함으로써, 제안하는 모델은 통신 환경 변화에 실시간으로 적응하여 자원 효율성을 높일 수 있다.

추후 연구로는 제안된 기법을 NS-3와 같은 네트워크 시뮬레이터에 구현하여 성능을 검증하고, 실제 차량 통신 환경과 유사한 시나리오에서 PDR, 지연 시간, 리소스 활용률 등의 지표를 종합적으로 평가할 계획이다.

ACKNOWLEDGMENT

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF-RS-2018-NR031059 and this work was supported by the Commercialization Promotion Agency for R&D Outcomes(COMPA) grant funded by the Korean Government(Ministry of Science and ICT). (RS-2023-00304695)

참 고 문 헌

- [1] Technical Specification Group Radio Access Network, "Study on NR Vehicle-to-Everything (V2X)" 3GPP, Tech. Rep. TR 38.885, Mar. 2019.
- [2] S. Jung, H.-R. Cheon, and J.-H. Kim, "Reducing consecutive collisions in sensing based semi persistent scheduling for cellular V2X," in Proc. IEEE Vehicular Technology Conference, Honolulu, Hawaii, USA, September 2019, pp. 1-5.
- [3] Hado van Hasselt, "Double Q-learning", Advances in Neural Information Processing Systems 23, 2011