

Cellular V2X에서 DQN 기반의 분산 혼잡 제어

이은화, 정구선, 최주영*, 문철

한국교통대학교, *연세대학교

20028606@ut.ac.kr, rntjs300@ut.ac.kr, cgy0103@yonsei.ac.kr, chmun@ut.ac.kr

Distributed Congestion Control Based on Deep Q-Network on in Cellular V2X

Eun-Hwa Lee, Gu-Sun Joung, Joo-Young Choi*, Cheol Mun

Korea National University of Transportation, *Yonsei University

요약

본 논문은 혼잡한 도로상황에 따른 C-V2X(cellular vehicle-to-everything) 통신 시스템에서 패킷 충돌로 인한 성능 열화를 극복하기 위해 DQN(deep Q-network)을 통해 각 차량마다 최적의 패킷 TTI(transmission time interval)를 조절하는 DQN 기반 분산 혼잡 제어 알고리즘을 제안한다. 모의실험을 통해 제안된 기법이 다양한 혼잡도 상황에서 PDR(packet delivery ratio)과 IPG(inter-packet gap) 측면에서 기존 알고리즘보다 성능을 향상시킬 수 있음을 확인하였다.

I. 서론

3GPP는 도로 안전과 트래픽의 효율적인 서비스를 지원하기 위해 Release 14에서 LTE(long term evolution) 기반의 C-V2X(cellular vehicle-to-everything)를 표준화했다. C-V2X mode 4에서 각 VUE(vehicular user equipment)는 셀룰러 인프라 지원 없이 SB-SPS(sensing-based semi-persistent scheduling)을 통해 독립적으로 무선자원을 할당하고 패킷을 전송한다. 하지만, 혼잡한 도로 상황에서는 다수의 VUE가 동일한 무선자원을 예약함으로써 송신 패킷들이 충돌하는 경향이 발생한다. 이러한 문제는 분산혼잡제어(distributed congestion control, DCC)를 통해 각 VUE가 차량 혼잡도에 따라 패킷 충돌을 억제할 수 있는 최적의 TTI(transmission time interval)와 전송전력으로 패킷을 송신함으로써 해결할 수 있다[1].

DCC는 SAE(society of automotive engineers) J2945/1[2]은 DSRC(dedicated short range communication)를 위한 TTI와 전송전력 제어 알고리즘을 표준화하였다. 반면에 C-V2X를 위한 DCC 표준으로써, DSRC용 DCC 표준인 SAE J2945/1을 C-V2X 표준에 적용한 SAE J3161/1[3]을 개발 중이다. 또한, C-V2X 표준에 SAE J2945/1 DCC 알고리즘을 적용한 연구들이 진행되어왔다[4]. 그러나, 다양한 혼잡도를 가지는 도로 상황에서 요구되는 성능을 만족시키지 못하고 있다.

따라서, 본 논문에서는 C-V2X mode 4 환경에서 다양한 혼잡도를 가지는 도로 상황에서 요구되는 성능을 만족시킬 수 있는 최적의 TTI를 찾는 DQN(deep Q-network) 기반 DCC 알고리즘을 제안하고 이의 성능을 분석한다.

II. SAE J2945/1 DCC 알고리즘

SAE J2945/1에서는 DSRC 표준에서 사용할 수 있는 DCC를 위한 TTI 제어 알고리즘을 제안하고 있다[2]. 시간 t 에서 패킷을 전송할 VUE i 는 패킷의 다음 전송 주기 TTI_i^t 를 이웃 차량 밀도(VUE density, VD_i^t)를 측정하고 이에 기초하여 결정한다. VD_i^t 는 VUE i 가 오류 없이 수신한 주변 차량들의 CAM(Cooperative Awareness Message) 메시지에 담긴 각 차량의 위치, 방향, 속도 정보 등을 이용하여 결정한다. VD_i^t 는 VUE i 의 반경 100 m 내에 위치한 이웃 VUE의 수로써 결정된다. VD_i^t 에 따른 TTI_i^t 는 다음과 같이 결정된다.

$$TTI_i^t = \begin{cases} 100 \text{ ms} & VD_i^t \leq B \\ (100 \cdot VD_i^t)/B \text{ ms} & B < VD_i^t < 6B \\ 600 \text{ ms} & 6B \leq VD_i^t \end{cases} \quad (1)$$

여기서 B 는 밀도 계수이며, SAE J2945/1에서는 25로 설정하고 있다. 이 경우, VD_i^t 가 25이하인 경우에는 기본 TTI 100ms를 사용하고, 25 초과 150 미만 경우에는 $4VD_i^t$ ms를 사용하고, 150 이상일 경우에는 최대 TTI인 600ms를 사용한다. 따라서, 밀도 계수 B 는 TTI를 결정하여 경쟁 VUE의 수를 조절하는 DCC의 성능을 결정하는 핵심 파라미터이다.

III. 제안하는 C-V2X를 위한 DQN 기반 DCC 알고리즘

SAE J2945/1에서는 TTI를 결정하기 위해 사용하는 밀도 계수 B 를 25라는 상수로 설정하고 있으나, 본 논문에서는 밀도 계수 B 가 도로의 혼잡도에 따라 최적화되어야 할 변수로써 각 VUE들이 각자 처한 도로의 혼잡도에 따라 최적의 B 를 찾는 알고리즘을 제안한다. 최적의 B 는 패킷 충돌을 최대한 억제하여 PDR(packet delivery ratio)을 최대화도록 결정되어야 한다. 다만, PDR 최대화만을 고려할 경우 TTI를 과도하게 증가시켜 CBR(channel busy ratio)을 과도하게 낮추어 전송 효율을 감소시킬 우려가 있다. 여기서 CBR은 VUE i 의 주변 다른 다수의 VUE에 의해 사용된 채널 이용률이다[5]. 따라서, 채널 사용률 CBR이 목표치에 근접하면서, PDR을 최대화도록 하는 최적화 문제를 다음과 같이 정의한다.

$$\begin{aligned} \text{P1: } \max_{TTI_i^t} & PDR_i[t, j] \\ \text{s.t. } \min_{TTI_i^t} & |CBR_i^{target} - CBR_i[t]| \end{aligned} \quad (2)$$

여기서 CBR_i^{target} 은 3GPP working document[5]와 같이 0.65로 설정하며, $CBR_i[t]$ 는 서브프레임 t 에서 측정된 VUE i 의 CBR이다. 그리고 서브프레임 t 의 j 번째 서브채널을 사용한 VUE i 의 패킷 성공률 $PDR_i[t, j]$ 는 다음과 같다.

$$PDR_i[t, j] = \sum_{k \in \delta} C_{i,k}[t, j] / \sum_{k \in \psi} F_{i,k}[t, j] \quad (3)$$

여기서 $F_{i,k}[t, j]$ 는 VUE i 에서 전송된 패킷을 나타내며 $C_{i,k}[t, j]$ 는 그중에서 오류 없이 전송된 패킷을 나타낸다. 집합 δ 는 수신에 성공한 VUE 집합으로 수신 VUE 집합 ψ 의 부분 집합이다. 경쟁 VUE의 수를 조절하는 밀도 계수 B 는 집합 δ 와 ψ 를 결정하고 이에 의해 PDR 성능이 결정된다.

본 논문에서는 C-V2X mode 4 환경에서 P1 문제를 DQN을 통해 해결하고자 한다. Q-learning은 MDP(Markov decision process)에서 agent가 의사 결정을 통해 얻은 state-action 쌍에 대해 $Q(s, a)$ 을 갖는다. $Q(s, a)$ 을 딥러닝을 통해 approximate하는 것을 DQN[6]이라고 한다. 제안하는 DQN 기반 DCC 알고리즘은

CTDE(centralized training decentralized execution) 기반으로 offline 상태에서 training을 할 때는 환경에서 패킷 성공 유무 정보를 공유하여 reward 수식 내 $PDR_i[t, j]$ 을 계산 및 학습을 수행하며, test를 할 때는 각 agent가 얻은 state 정보만으로 테스트를 수행한다. Agent는 시뮬레이션 환경에 있는 모든 VUE이며, 기본 파라미터들은 다음과 같다. 시간 t 에서 패킷을 전송하는 agent i 의 state는 $s_t^i = CBR_i[t]$ 이며, action은 $a_t^i = \{5, 25, 45, 65\}$, 즉 SAE J2945/1 TTI 제어의 파라미터인 B 이며, 그리고 reward는 $r_t^i = PDR_i[t, j] - |0.65 - CBR_i[t]|$ 이다.

그림 1은 제안하는 DQN 기반 DCC 학습과정을 나타내며, t 에서 패킷을 전송하는 VUE i , 즉 agent i 의 학습과정을 나타낸다. VUE의 network 구조는 100개의 노드를 가진 fully connected layer 2개, ReLU(rectified linear unit) layer 2개, 그리고 a_t^i 내 개수만큼 노드를 가지는 output으로 구성된다. 또한, decaying ϵ -greedy의 초기 값은 1, learning rate는 0.01, discount factor는 0.99, mini batch 크기는 256, 그리고 replay memory 크기는 100,000이다. agent i 는 s_t^i 을 input으로 main network에서 output, $Q(s, a; \theta)$ 에 ϵ -greedy policy를 적용하여 a_t^i 을 구하고, 이를 (1)에 대입하여 TTI_t^i 을 구한다. 그리고 다음 time step에서 r_t^i 및 s_{t+1}^i 을 관찰하고 $(s_t^i, a_t^i, r_t^i, s_{t+1}^i)$ 을 replay memory에 쌓는다. replay memory의 mini batch 크기의 transition data를 통해 $L(\theta_t)$ 을 다음과 같이 구한다.

$$L(\theta_t) = \mathbb{E}[(r(s, a) + \gamma \max_{a'} \hat{Q}(s', a'; \theta^-) - Q(s, a; \theta))^2] \quad (4)$$

(4)를 통해 시간 t 에서 패킷을 전송하는 agent i 의 DQN network는 target network에서 $\hat{Q}(s', a'; \theta^-)$ 와 main network에서 $Q(s, a; \theta)$ 의 MSE(mean square error)를 줄여나가는 방향으로 gradient를 구하고, main network 파라미터 θ 를 업데이트하며 학습한다. agent i 의 target network의 파라미터 θ^- 는 100개의 time step 간격으로 agent i 의 main network 파라미터 θ 와 동기화한다.

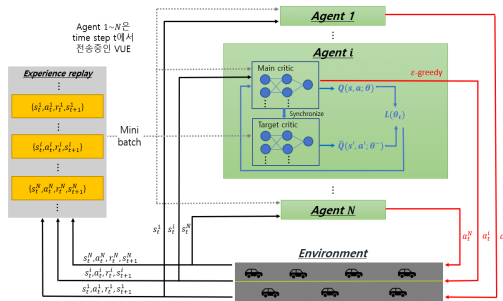


그림 1. 제안하는 DQN 기반 DCC 학습과정

IV. 모의실험 결과

시뮬레이션은 C-V2X 네트워크의 자원할당을 평가하기 위한 동적 시뮬레이터인 LTEV2Vsim[7]을 기반으로 수행된다. 시뮬레이션의 환경 및 파라미터는 표 1과 같다. 제안하는 알고리즘의 성능을 PDR과 평균 IPG(inter-packet gap) 관점에서 분석한다. 여기서 IPG는 동일한 송신 VUE와 수신 VUE 사이에서 연속적으로 오류 없이 수신된 2개의 패킷 사이 시간 간격이며, 평균 IPG는 송수신기 거리에 따른 시뮬레이션 환경 내 모든 VUE의 IPG에 대한 평균이다.

표 1. 시뮬레이션 환경 및 시스템 파라미터

표준 버전	C-V2X mode 4	MCS index	6(=QPSK, 0.48)
시뮬레이션시간	20초	패킷 크기	300 bytes
중심 주파수	5.9 GHz	SPS 임계값	-94 dBm
대역폭	10 MHz	전파모델	WINNER II, B1
도로 환경	2km 고속도로, 왕복 8 차선		
차량 속도	60km/h, 표준편차 3km/h		
차량 밀도	2km 내 400, 500대		

그림 2는 2km 내 400대와 500대의 VUE 밀도에서 제안하는 알고리즘과 기존 SAE J2945/1 TTI 제어 알고리즘에서 파라미터 $B=25$ 일 때, PDR과 IPG 성능을 나타낸다. 400대 혼잡도 상황에서 제안하는 알고리즘은 기존 SAE J2945/1 알고리즘에 비해 2.9~5.3% PDR 이득을 가지며, 3~12 ms 더 작은 평균 IPG를 제공한다. 500대 혼잡도 상황에서 제안하는 알고리즘은 0~2.4% PDR 이득을 가지며, 46~55 ms 더 작은 평균 IPG를 제공한다. 400대에서 500대로 혼잡도가 증가할 때, 제안하는 알고리즘의 평균 IPG가 12~28 ms 증가에 그친 반면에 기존 SAE J2945/1 알고리즘은 65~73 ms의 더 높은 증가를 보여준다. 동시에 제안하는 알고리즘은 혼잡도 증가에 따라 PDR이 감소하지만 기존 SAE J2945/1 알고리즘보다 우수한 PDR 성능을 제공한다. 따라서, 제안하는 알고리즘은 다양한 혼잡한 상황에서 PDR과 IPG 측면에서 기존 SAE J2945/1 알고리즘에 비해 성능이 향상됨을 보여준다.

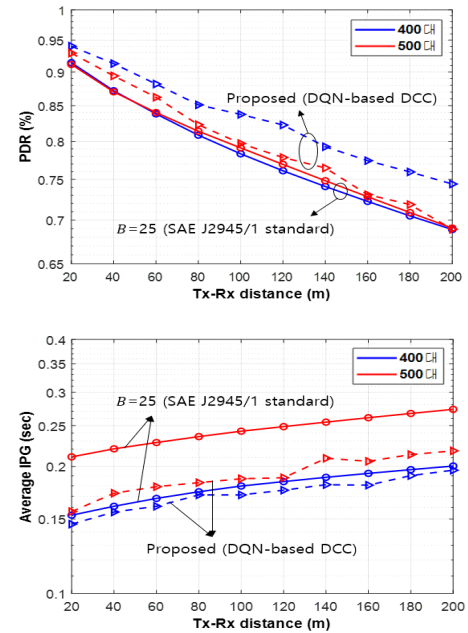


그림 2. 다양한 혼잡도에서 PDR과 IPG 성능 비교

ACKNOWLEDGMENT

본 연구는 연구재단 기초연구지원사업(2018R1D1A3B0705032714)과 과학기술정보통신부 및 정보통신기획평가원의 Grand ICT 연구센터지원사업(IITP-2021-0-01462)의 연구결과로 수행되었음.

참고 문헌

- [1] 정구선, 정창규, 문철, "C-V2X를 위한 지역 혼잡도 정보를 이용한 분산 혼잡 제어," 한국전자학회논문지, vol. 32, no. 2, pp. 101-109, Feb. 2021.
- [2] SAE International, on-board system requirements for V2V safety communications, Standard Doc. J2945/1, 2016.
- [3] SAE International, on-board system requirements for LTE-V2V safety communications(under development), Standard Doc. J3161/1, Aug. 2019.
- [4] B. Toghi, M. Saifuddin, Y. P. Fallah, and M. O. Mughal, "Analysis of distributed congestion control in cellular vehicle-to-everything networks," in 2019 IEEE 90th Veh. Technol. Conf.(VTC-Fall), 2019.
- [5] Qualcomm, "R1-1611594. congestion control for V2V," 3GPP TSG RAN1-87, Tech. Rep., Nov. 2016.
- [6] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," Nature, vol. 518, no. 7540, pp. 529-533, Feb. 2015.
- [7] Web page of LTEV2Vsim. [Online]. Available: <https://github.com/alessandrobazzi/LTEV2Vsim>