

# 전파 환경을 고려한 심층강화학습 기반 QoS 적응형 분산 혼잡 제어

전병철, 남태식, 조한신\*

한밭대학교, 연세대학교, \*한양대학교

30221177@edu.hanbat.ac.kr, ts.nam@yonsei.ac.kr \*hsjo23@hanyang.ac.kr

## QoS-Adaptive Distributed Congestion Control Based on Deep Reinforcement Learning in Consideration of Radio Environment

Jeon Byeong Cheol, Nam Tae Sik, Jo Han-Shin\*

Hanbat National Univ., Yonsei Univ., \*Hanyang Univ.

### 요약

분산 혼잡 제어 (Distributed Congestion Control, DCC)는 높은 밀도의 차량 네트워크에서 채널 혼잡을 완화하고, 통신 성능을 개선하는 기술이다. 기존의 DCC 기술은 Quality of Service (QoS) 요구사항을 고려하지 않고 채널 혼잡을 완화하는 방향으로 동작한다. 이러한 설계는 과도한 DCC 동작으로 인하여 다른 QoS를 저하시킬 수 있다. 이와 같은 문제를 해결하기 위해 심층강화학습 기반 QoS 적응형 DCC 알고리즘을 제안한다. 시뮬레이션은 두 가지 전파 환경을 고려한 도심 시나리오에서 평가하였으며, 시뮬레이션 결과 기존 DCC 기술보다 목표 QoS에 더 근접한 결과를 확인하였다.

### I. 서론

미국의 국제자동차기술자협회 (Society of Automotive Engineers, SAE)는 자율주행의 자동화 수준에 따라 6단계로 분류하고 있다 [1]. 완전 자율주행 수준인 레벨 4는 운전자의 개입 없이 서비스를 지원해야 하며, 자동차의 전장 부품만으로 해당 서비스를 지원하기에는 한계가 있다. 이러한 한계를 극복하기 위해 차세대 지능형 교통 시스템 (Cooperative-Intelligent Transport Systems, C-ITS)와의 통합이 중요하다. C-ITS는 주행 중 차량이 운전자에게 주변의 교통상황, 전방의 장애물 등 사고 위험 정보를 실시간으로 제공한다. 이와 같은 C-ITS를 구축하기 위한 필수 기술로는 V2X (Vehicle-to-Everything)이 있다.

이동 통신 표준화 기관인 3GPP (3rd Generation Partnership Project)는 LTE 통신 프로토콜 기반의 C-V2X (Cellular-V2X)를 표준화하였으며, C-V2X는 CAM (Cooperative Awareness Message)를 통해 차량의 위치, 속도 그리고 방향 등의 정보를 주기적으로 주변 차량과 정보를 교환한다.

차량 통신 네트워크에서 차량 밀도의 증가는 한정된 무선 자원을 다수의 차량이 점유하기 때문에 패킷 충돌, 누적 간섭 증가 등의 문제를 야기하며 통신 성능의 저하를 일으킨다. 이러한 고밀도 환경의 통신 성능 개선을 위한 대표적인 기술로는 분산 혼잡 제어 (Distributed Congestion Control, DCC)가 있다. DCC는 패킷의 전송 주기, 전송 전력, 전송 MCS (Modulation and Coding Scheme) 등을 조절하여 채널 혼잡도를 낮추고, 통신 성능을 개선한다. 하지만 기존 차량 통신 관련 표준화 기관에서 제안하는 DCC 기술은 채널 혼잡도를 감소시키는 것에 초점을 맞추어 통신 성능을 개선하기 때문에 서비스의 QoS 요구사항을 만족할 수 있는 환경이 제한적이다.

따라서, 본 논문에서는 심층강화학습 기반의 C-V2X의 V2V 시나리오 QoS 요구사항에 적응적인 DCC 알고리즘을 제안한다. 심층강화학습은 강화학습과 딥러닝을 결합한 기계 학습의 하위 분야로 V2X 분야에도 다양하게 연구되고 있다. 또한 도심 환경의 전파 특성을 고려한 학습 모델 설

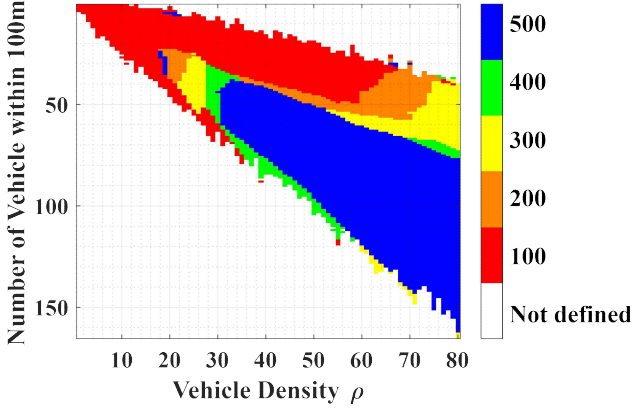
계를 통하여 실제 C-V2X 시스템 구축을 위한 알고리즘의 가능성을 검증한다.

### II. 최적화 문제 정의

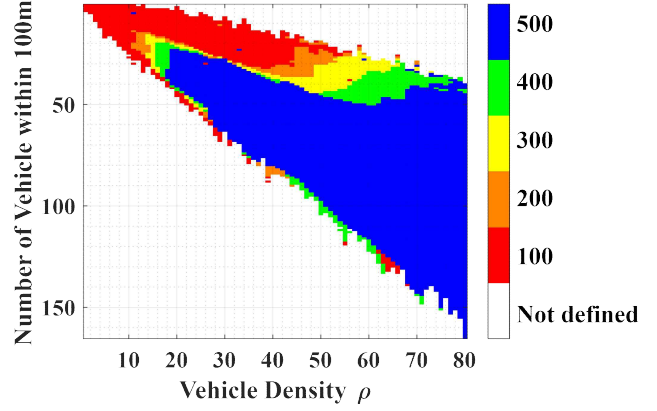
차량은 패킷 전송 주기를 통해 차량 네트워크의 혼잡도를 낮추어 PDR (Packet Delivery Ratio)를 향상시킬 수 있다. 하지만 혼잡 수준에 대한 과도한 전송 주기 제어의 동작은 지연 시간을 증가시킬 수 있다. 따라서 혼잡한 상황에서 적절한 수준의 메시지 전송 주기 제어 동작은 필수적이다. 따라서 본 연구에서 차량은 목표 통신 범위 내에서 목표 QoS 즉, PDR을 만족하는 수준까지만 전송 주기 제어를 동작시킨다. 최적화 문제는 다음과 같이 정의된다.

$$\begin{aligned} P1 : \min_{\mathbf{U}} & \left| \psi_{\rho, \beta}^* - \psi_{\rho, \beta}^{avg} \right|, \\ \psi_{\rho, \beta}^{avg} &= \frac{\sum_{v \in \mathbf{V}} \sum_{z \in \mathbf{Z}} \sum_{n \in \mathbf{N}} \delta[n, c, \beta]}{\sum_{v \in \mathbf{V}} \sum_{z \in \mathbf{Z}} \sum_{n \in \mathbf{N}} \lambda[n, c, \beta]}, \\ s.t. C1 : \mathbf{U} &= \{u_1, u_2, u_3, \dots, u_M\} \\ u_m &\in \{\delta_1, \delta_2, \delta_3, \dots, \delta_F\} \end{aligned} \quad (1)$$

식 (1)에서  $\psi_{\rho, \beta}^*$ 는 차량 밀도  $\rho$ 에 대해 목표 통신 범위  $\beta$ 에서의 차량 네트워크 목표 PDR이며,  $\psi_{\rho, \beta}^{avg}$ 는 차량 네트워크 평균 PDR이다.  $u_m$ 은 패킷 생성 시간 간격의 상한값을 의미하며, 이는  $\delta$  집합의 원소이다. 패킷 생성 시간은 최소 100 ms와  $u_m$  사이의 무작위 정수로 결정된다.  $\mathbf{U}$ 는 시나리오 내 모든 차량의 패킷 생성 시간 간격의 집합이다. 결과적으로 P1은 차량 네트워크 상 모든 VUE의 패킷 생성 시간을 분산적으로 조절하여 통신 범위  $\beta$  내 평균 PDR과 목표 PDR의 차이를 최소화하는 문제로 정의된다.



(a)  $\psi_{\rho,\beta}^* = 80\%$



(b)  $\psi_{\rho,\beta}^* = 90\%$

그림 1. 학습된 전송 주기 정책

### III. 심층강화학습 모델

멀티 에이전트 심층강화학습은 다수의 차량이 공존하는 V2V 시나리오에서 효과적인 학습 프레임워크이다. 또한 V2V의 반 이중 통신(Half-Duplex)로 인해 전송하는 시간의 무선 채널 정보를 알 수 없다. 이러한 특성을 고려하여 CTDE (Centralized Training and Distributed Execution) 구조를 기반으로 학습을 수행한다. 해당 학습 모델에서 중앙 집중식 컨트롤러는 RSU (Roadside Unit)으로 대체할 수 있으며, 학습을 위한 상태는 다음과 같다.

$$s_t = [\eta_t^{(m)}, \rho] \quad (2)$$

여기서  $\eta_t^{(m)}$ 는 시간  $t$ 에서  $m$ 번째 차량으로부터 측정된 100 m 이내 차량 수이며,  $\rho$ 는 차량 밀도이다. 컨트롤러는 상태 정보를 기반으로 목표 QoS에 근접하도록 전송 주기를 조절해야 한다. 컨트롤러의 행동 공간은 다음과 같다.

$$A = \{\delta_1, \delta_2, \delta_3, \dots, \delta_F\}, a_t \in A \quad (3)$$

여기서  $F=5$ 이며,  $a_t = \{100, 200, 300, 400, 500\}$ 으로 설정된다. 학습의 목표인  $\psi_{\rho,\beta}^*$ 와  $\psi_{\rho,\beta}^{avg}$  사이를 최소화하기 위한 학습을 진행한다. 각 차량은 전역 정보인  $\psi_{\rho,\beta}^{avg}$ 를 알 수 없지만, 중앙 컨트롤러는 전역 정보를 안다고 가정한다. 학습의 보상 설계는 다음과 같다.

$$r_t = -|\psi_{\rho,\beta}^* - \psi_{\rho,\beta}^{avg}| \quad (4)$$

학습이 종료된 뒤, 각 차량은 학습된 정책을 컨트롤러로부터 공유받으며 실행 단계에서 컨트롤러의 개입은 요구되지 않는다. 실행 단계에서의 각 차량은 로컬 정보  $\eta_t^{(m)}$ 만을 이용하여 독립적으로 동작한다.

### III. 시뮬레이션

제안하는 알고리즘을 평가하기 위한 시뮬레이션 파라미터는 표 1과 같다. 시뮬레이션 환경은 두 가지 전파 환경 LOS (Line-of-Sight) 및 NLOS (Non-Line-of-Sight)가 공존하는 도심 환경에서 수행되었다. 또한 목표 QoS는 통신 범위  $\beta$ 와 PDR로 채택했으며, 각각  $\beta=100$  m 와  $\psi_{\rho,\beta}^*=80, 90\%$ 이다.

표 1. 시뮬레이션 파라미터

파라미터	값
시뮬레이션 시간 [s]	30
시뮬레이션 환경	Urban Grid
차량 수 [veh]	19-1615
MCS index	3
채널 대역폭 [MHz]	10
채널 모델	TR 37.885
안테나 이득 [dB]	3
목표 통신 범위 $\beta$ [m]	100
목표 PDR $\psi_{\rho,\beta}^*$ [%]	80, 90

그림 1은 각각  $\psi_{\rho,\beta}^*=80, 90\%$  일 때 학습된 전송 주기 정책을 보여준다. (a)와 (b) 모두 100 m 이내 차량의 수가 증가하고, 차량 밀도가 증가할수록 즉, 무선 채널 환경이 혼잡할수록 더 높은 행동을 선택한다. (b)는 (a)보다 더 높은 목표 QoS로 인해 혼잡도가 낮은 환경에서 더 높은 행동을 취한다.

### III. 결론

본 연구는 전파 환경을 고려한 심층강화학습 기반의 QoS 요구사항에 적응적인 전송 DCC 알고리즘을 제안한다. 제안하는 알고리즘은 기존 DCC 기술에서 고려하지 않았던 QoS 요구사항을 활용함으로써 목표 QoS 수준을 효과적으로 만족시켰다. 이와 같은 과정을 통하여 C-ITS를 구축하기 위한 C-V2X 시스템의 실제 적용 가능성을 확인했다. 향후 연구에서는 제안하는 알고리즘을 시간에 따라 환경이 급변하는 상황에 적용 가능한 보다 일반화된 심층강화학습 모델로 확장하는 것을 고려한다.

### ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2022-0-01053, 다중 통신기술 네트워크 로드밸런싱 기술개발)

### 참고 문헌

- [1] (R) *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, Standard SAE J3016, Apr. 2021.