

다중 에이전트 강화학습을 이용한 RSU 선택 및 AoI 최적화

허주희, 김수빈, 박수현*

숙명여자대학교

heojuhui@sookmyung.ac.kr, rlatnqls0822@sookmyung.ac.kr, *soohyun.park@sookmyung.ac.kr

MARL-Based Optimization of AoI-Aware RSU Selection

Juhui Heo, Subin Kim, Soohyun Park*

Sookmyung Women's Univ.

요약

자율 주행 차량은 신뢰성·에너지 효율·고속 교통을 위해 정밀한 센서 데이터에 의존하지만, 센서 이상은 신뢰성을 저하시켜 에너지 소비를 증가시킨다. 본 논문은 최적 RSU 전송 정책과 오류 없는 통신, 낮은 AoI를 통한 정보 최신성 유지, 그리고 균등한 RSU 활용을 학습하는 CommNet 기반 심층 강화학습(Deep Reinforcement Learning, DRL) 프레임워크를 제안하며, 이는 부분 관측의 자율 주행 차량 플레톤과 같은 다중 에이전트 환경에 적합하다. 독립적 다중 에이전트 강화학습(Independent Multi-Agent Reinforcement Learning, Independent MARL), 심층 Q-네트워크(Deep Q-Network, DQN), 탐욕 정책(greedy policy)과의 비교 실험에서 제안 기법은 0% 오류 전송, 최대 63.2%의 AoI 감소, 그리고 가장 효율적인 통신 비용(0.034%)을 달성하여, 안전한 협력 제어와 신뢰성, 정보 최신성, 에너지 효율 향상을 입증한다.

I. 서론

자율 주행 기술의 발전으로 빠르고 신뢰성 있는 차량-인프라(vehicle-to-infrastructure, V2I) 통신에 대한 수요가 증가하고 있다. 자율 주행 차량(Autonomous Vehicles, AVs)에서 노변 장치(Roadside units, RSUs)로 전송되는 오류 없는 실시간 데이터는 중복 전송과 에너지 소비를 줄이고, 정확한 상황 인식과 신속한 이상 탐지를 가능하게 한다 [1].

밀집 교통 환경에서 신뢰성 있는 RSU 통신을 위해서는 데이터 신뢰성, 정보 신선도(Age of Information, AoI), 그리고 균등한 RSU 활용이 필수적인 성능 지표이다. 오류 없는 전송은 중복을 감소시키고 정확도를 향상시키며, AoI는 시의적절한 제어 결정을 지원한다. 또한 RSU 부하의 균형은 병목과 재전송을 완화하여 에너지 효율을 향상시킨다 [2].

그러나 다수의 차량을 대상으로 신뢰성과 정보 신선도를 동시에 만족하는 RSU 통신은 여전히 도전적인 문제로 남아 있으며, 이는 차량 간 간격이 짧고 실시간 정보 공유에 의존하는 플레톤 환경에서 더욱 두드러진다. RSU는 안전한 차량 협조와 도로 수용량 및 에너지 효율 향상에 기여하지만, 기존 하드웨어·알고리즘·AI 기반 접근법은 높은 비용, 모델링 복잡성, 그리고 협력 인지 능력의 한계를 가진다.

이러한 한계를 극복하기 위해 본 논문은 CommNet 기반 DRL 프레임워크를 제안한다(그림1). 공유된 잠재 통신 표현을 통해 협력적 정책 최적화를 가능하게 하고 센서 잡음을 완화함으로써, 통신 신뢰성, 정보 신선도, 그리고 에너지 효율성을 플레톤 환경에서 효과적으로 향상시킨다 [3].

II. 시스템 모델

본 연구에서는 각 차량의 정보 신선도를 AoI로 모델링한다. 시간 슬롯 t 에서 차량 v_n 의 AoI를 $G_n(t)$ 로 정의하며, 전송 실패 또는 RSU 미선택 시 AoI는 $G_n(t+1) = G_n(t) + 1$ 로 증가하며, 차량 v_n 이 직접 전송하거나 플레톤 리더에게 위임해 RSU u_m 로 전송에 성공한 경우 $G_n(t+1) = 0$ 으로 초기화된다. 네트워크는 각 방향 3.5m 폭의 3차선을 갖는 길이 1000m 순환 도로에 차량 집합 $V = \{v_1, \dots, v_N\}$ 과 RSU

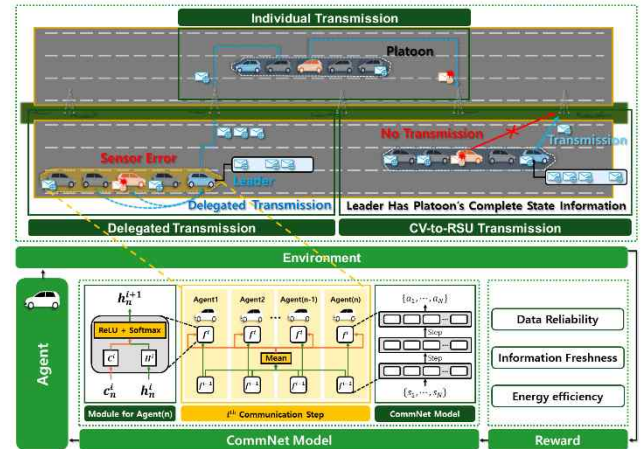


그림 1 차량 플레톤 시나리오를 위해 제안된 CommNet 기반 협력적 심층 강화학습 프레임워크의 전체 구조

집합 $U = \{u_1, \dots, u_N\}$ 가 배치된 구조로 가정한다. 차량들은 고정 크기의 플레톤을 형성하며, 선두 차량은 리더로서 멤버 차량의 데이터를 집계해 RSU와 통신한다. 각 차량은 매 시간 슬롯마다 전송 여부를 독립적으로 결정하고, 플레톤 내 차량 간 통신(Vehicle-to-Vehicle, V2V)은 신뢰적으로 이루어진다고 가정하며, RSU로의 전송 시 에너지가 소모된다.

III. 제안 알고리즘

본 연구에서는 CommNet 기반 DRL 프레임워크를 채택한다. 각 차량은 지역 관측 정보를 잠재 표현으로 인코딩하고, 공유된 통신 계층을 통해 특정 정보를 교환한 후, 전송하지 않는 선택지를 포함하는 RSU 선택 행동 확률 벡터를 생성한다. 통신 가능 범위 250m를 초과하는 RSU 선택지는 행동 선택 이전에 마스킹되며, 크리티크는 공동 관측 정보를 기반으로 CTDE 구조에서 학습된다.

시간 t 에서 중앙 집중 크리티크는 차량의 정보 신선도, 차량과 RSU 간 거리, RSU의 수용량 상태, 센서 오류 상태, 그리고 플레톤 리더 정보를 포함

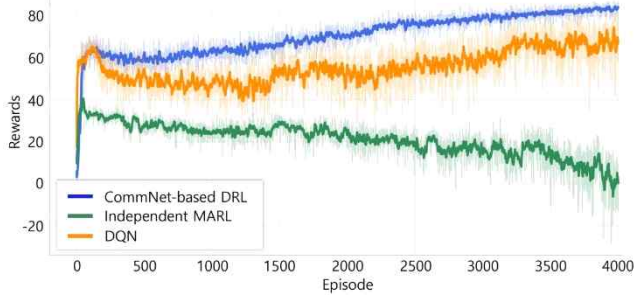


그림2 에피소드 별 보상 수렴 비교

하는 전역 시스템 상태를 입력으로 사용한다. 각 차량은 다음과 같이 시간 슬롯당 최대 하나의 RSU를 선택한다. 또한 플레톤 리더가 전송을 수행하는 경우, 위임이 적용되어 멤버 차량들의 전송은 리더의 행동으로 통합된다. 보상 함수는 개인적 요소와 협력적 요소로 구성된다. 개인적 요소는 센서 오류 데이터의 직접 전송, 거리 기반 통신 비용, 그리고 거리로 인한 전송 손실을 반영하여 불필요하거나 비효율적인 전송을 억제한다. 반면, 협력적 요소는 RSU 활용의 균형과 정보 신선도 향상을 고려하여, 플레톤 전체 관점에서 효율적이고 시의적절한 데이터 전송을 유도한다. 보상 함수는 각 요소의 중요도에 따라 가중치가 부여되어 최종 보상에 반영된다.

IV. 성능 평가

A. 실험 설정

제안된 CommNet 기반 DRL 프레임워크의 성능을 평가하기 위해 세 가지 베이스라인을 고려한다. Independent MARL은 CTDE 프레임워크 하에서 에이전트 간 통신 없이 각 AV가 지역 정보만을 이용해 독립적으로 의사 결정을 수행한다. DQN은 각 AV가 지역 관측값을 이산적 전송 행동으로 매핑하는 독립적으로 학습된 DQN 액터를 실행하며, 유효하지 않은 행동은 Q값 최대화 이전에 마스킹되고 에이전트 간 통신은 사용하지 않는다. Greedy에서는 시간마다 각 AV가 가장 인접한 RSU를 선택한다.

B. 평가 결과

그림2는 CommNet 기반 DRL이 가장 높고 안정적인 보상 수렴을 달성함을 보여준다. 이는 공유 통신을 통해 오류 전송을 효과적으로 억제하고 균등한 RSU 선택을 유도하기 때문이다. 반면, DQN은 분산형 Q-러닝의 비정상성 문제로 인해 뚜렷한 진동과 제한적인 수렴 성능을 보인다. Independent MARL은 에이전트 간 통신 부재로 협력적 행동과 리더 위임 전략의 형성이 어렵고, 그 결과 비체계적인 전송 결정, RSU 활용 불균형, AoI 증가로 인해 보상 향상이 제한된다.

표1은 각 알고리즘의 성능을 비교한 결과를 요약한다. Greedy는 매 단계 전송과 리더 위임을 통해 직접 오류 전송은 회피하지만, 차량이 인접 RSU에 집중되어 혼잡이 발생하며, 이로 인해 전송 성공률 저하, AoI 증가, 그리고 RSU 분포 엔트로피의 급격한 감소가 나타난다. Independent MARL은 에이전트 간 통신 부재로 협력적 RSU 선택이 어려워 전송 불균형과 높은 AoI가 지속되며, 직접 오류 전송 또한 완전히 제거하지 못한다. 통신 없이 독립적으로 학습된 DQN은 비교적 높은 RSU 분포 엔트로피와 전송률을 달성하지만, 협력 부재로 인해 행동이 불안정해지고 직접 오류 전송이 빈번히 발생하여 RSU 활용 효율과 AoI 측면에서 한계를 보인다. 반면, CommNet 기반 DRL은 결합 데이터를 플레톤 리더에게 일관되게 위임함으로써 직접 오류 전송을 제거하고, 균등한 RSU 활용 하에서 가장 낮은 통신 비용과 최소 평균 AoI를 달성했다.

표1 평가 대상 알고리즘 간 성능 비교

알고리즘	RSU 분포 엔트로피	통신 비용	평균AoI (time step)	전송률 (%)	직접오류 전송비율 (%)	위임오류 전송비율 (%)
CommNet based DRL	2.005	0.034	0.15	18.87	0.000	0.314
Independent MARL	0.768	0.022	73.11	14.46	1.199	0.000
DQN	2.278	0.034	0.41	21.39	3.036	0.317
Greedy	0.000	0.041	48.20	13.33	0.000	0.222

V. 결론

본 연구는 차량 네트워크 환경에서 오류 없는 전송, 향상된 정보 신선도, 그리고 균등한 RSU 활용을 달성하는 CommNet 기반 DRL 프레임워크를 제안한다. 공유된 잠재 통신 표현을 통해 안정적인 협력 정책을 학습하고 오류 전송을 완전히 제거하였으며, 실험 결과 전송률 대비 가장 효율적인 통신 비용과 모든 베이스라인 중 최소 AoI를 달성하였다. 이러한 결과는 통신 기반 협조 제어가 신뢰성과 에너지 효율을 갖춘 플레투닝에 효과적인임을 입증하며, 향후 연구에서는 동적 플레톤 구성, 이질적 센서, 그리고 대규모 교통 환경으로의 확장을 고려할 수 있다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부 및 산업기술평가관리원(KEIT)의 지원(RS-2024-00442168) 및 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업(IITP-2026-RS-2024-00436887)의 연구결과로 수행되었음. 본 논문의 교신저자는 박수현임.

참고 문헌

- [1] F. Matos, J. Bernardino, J. Durães, and J. Cunha, "A Survey on Sensor Failures in Autonomous Vehicles: Challenges and Solutions," *Sensors*, vol. 24, no. 16, p. 5108, August 2024.
- [2] K. Jyostna and G. S. Sharma, "Enhanced Sustainability in Green Vehicular Networks Through Energy-Aware Resource Management and Collaborative Computing," in *Proc. 2025 International Conference on Modern Sustainable Systems (CMSS)*, Shah Alam, Malaysia, October 2025, pp. 1090 - 1097.
- [3] T. -Y. Tung, S. Kobus, J. P. Roig and D. Gündüz, "Effective Communications: A Joint Learning and Communication Framework for Multi-Agent Reinforcement Learning Over Noisy Channels," in *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2590-2603, Aug. 2021.
- [4] W. Pi, P. Yang, D. Duan, C. Chen, X. Cheng, L. Yang, and H. Li, "Malicious user detection for cooperative mobility tracking in autonomous driving," *IEEE internet of things journal*, vol. 7, no. 6, pp. 4922 - 4936, February 2020.
- [5] W. Xu, C. Yan, W. Jia, X. Ji, and J. Liu, "Analyzing and enhancing the security of ultrasonic sensors for autonomous vehicles," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 5015 - 5029, December 2018.
- [6] A. A. Al-Habob, H. Tabassum, and O. Waqar, "Non-orthogonal age optimal information dissemination in vehicular networks: A meta multi objective reinforcement learning approach," *IEEE Transactions on Mobile Computing*, vol. 23, no. 10, pp. 9789 - 9803, February 2024.