

안전 임계형 차량-사물 통신 기반 통합 센싱-통신 환경에서의 자원 할당을 위한 제약 인지형 심층 강화학습

엄수빈¹, 서은태², 황상원²
한국자동차연구원¹, 국립부경대학교²

sbeom@katech.re.kr, syt06100@naver.com, s.won.hwang@pknu.ac.kr

A Constraint-Aware Deep Reinforcement Learning for Resource Allocation in Safety-Critical Vehicle-to-Everything Integrated Sensing-and-Communication

Subin Eom¹, Euntae Seo², and Sangwon Hwang²

Korea Automotive Technology Institute¹, Pukyong National University²

Abstract

본 논문은 안전 임계형 V2X-ISAC 환경에서 도로변 기지국(roadside unit, RSU)이 교차로 사각지대의 은닉 객체를 센싱하면서 주변 차량에 긴급 경고 메시지를 전달하는 시나리오를 고려한다. 통합 센싱-통신(integrated sensing and communication, ISAC) 구조에서는 제한된 시간 자원을 센싱과 통신이 공유하므로, 센싱 비중을 높이면 추적 정확도는 향상되지만 통신 기회가 감소하고, 반대로 통신 비중을 높이면 경고 지연은 줄어들지만 센싱 불확실성이 증가할 수 있다. 특히 안전 임계형 V2X-ISAC 환경에서 센싱 조건은 완화 가능한 성능 목표가 아니라 반드시 만족되어야 하는 안전 임계 센싱 제약(safety-critical sensing constraint)이며, 이를 위반하면 은닉 객체의 상태 추정 오차가 증가하여 위험 상황 인지의 적시성이 저하될 수 있다. 이에 본 논문은 센싱 제약을 반드시 만족하면서 동시에 통신 지연을 최소화하는 자원 할당 문제를 정의하고, 이를 위한 제약 인지형(constraint-aware) 심층 강화학습(deep reinforcement learning, DRL) 기법을 제안한다.

i Introduction

차량-사물 통신(vehicle-to-everything, V2X) 환경에서는 주변 상황 인지와 위험 정보 전달이 동시에 요구된다. 최근 통합 센싱-통신(integrated sensing and communication, ISAC) 기반 차량 네트워크에서는 도로변 기지국(roadside unit, RSU)이 센싱과 통신을 공통 프레임 자원 위에서 함께 운용할 수 있으므로, 제한된 자원을 어떻게 분할할 것인가가 핵심 문제가 된다 [3, 4]. 특히 안전 임계형 V2X-ISAC 환경에서 센싱 조건은 반드시 만족되어야 하는 안전 임계 센싱 제약(safety-critical sensing constraint)이며, 이를 위반하면 은닉 객체의 상태 추정 오차와 추적 불확실성이 증가하여 위험 상황 인지와 경고 판단의 적시성이 저하될 수 있다. 반면 센싱 자원을 과도하게 늘리면 경고 전달 지연이 증가할 수 있으므로, 센싱 제약을 반드시 만족하면서 동시에 통신 지연을 최소화하는 자원 할당이 필요하다. 본 논문은 이러한 환경에서 통신-센싱 자원 할당 문제를 정의하고, 이를 해결하기 위한 제약 인지형(constraint-aware) 심층 강화학습(deep reinforcement learning, DRL) 기반 자원 제어 기법을 다룬다.

ii V2X-ISAC System Model

본 논문은 Fig. 1과 같이 RSU가 교차로 사각지대의 은닉 객체를 센싱하면서 동시에 주변 차량에 경고 메시지를 전송하는 안전 임계형 V2X-ISAC 시나리오를 고려한다. 시간은 슬롯 집합 $\mathcal{T} = \{1, \dots, T\}$ 와 슬롯 인덱스 $t \in \mathcal{T}$ 로 구분되며 각 슬롯의 길이는 τ 이다. RSU는 슬롯마다 센싱 비율 $a^{(t)} \in [0, 1]$ 를 선택하고, 슬롯 t 에서 센싱에 할당된 시간과 통신에 할당된 시간을 각각 $\tau_s^{(t)}$ 와 $\tau_c^{(t)}$ 로 둔다.

즉, 센싱 비율 $a^{(t)}$ 에 따라 전체 슬롯 시간이 센싱과 통신에 분할된다.

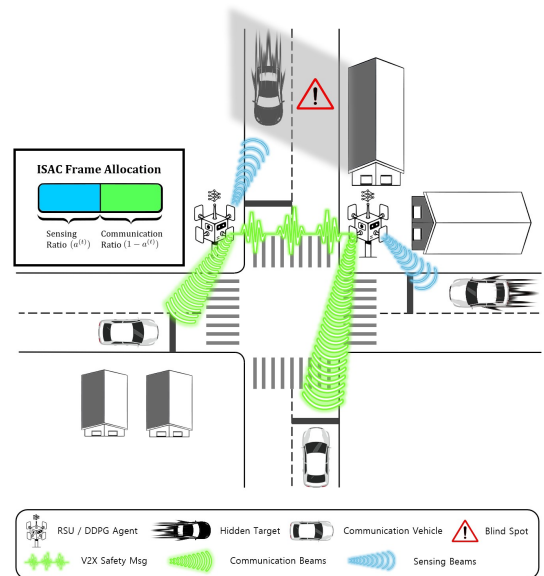


Figure 1: 교차로 사각지대 V2X-ISAC 시스템 모델

센싱 측에서는 비협력 은닉 객체를 확장 칼만 필터(extended Kalman filter, EKF)로 추적한다 [2]. 센싱 자원 할당은 측정 품질을 제어하며, 측정 오차 공분산은

$$\tau_s^{(t)} = a^{(t)}\tau, \quad \tau_c^{(t)} = (1 - a^{(t)})\tau \quad (1)$$

$$\mathbf{R}^{(t)} = s(a^{(t)})\mathbf{R}_0 \quad (2)$$

로 두며, 여기서 $\Delta_s = s_{\max} - s_{\min}$, $s(a^{(t)}) = \text{clip}(s_{\max} - \Delta_s a^{(t)}, s_{\min}, s_{\max})$, 그리고 $\text{clip}(x, a, b) = \min(\max(x, a), b)$ 이다. 따라서 $a^{(t)}$ 가 커질수록 $\mathbf{R}^{(t)}$ 가 작아지고, 더 정확한 EKF 측정 갱신이 이루어진다. 센싱 신뢰도는 standard EKF-based tracking uncertainty metric [2]에 따라 posterior position covariance의 trace로 정의한

$$e_{\text{post}}^{(t)} = \text{tr}(\mathbf{P}_p^{(t|t)}) \quad (3)$$

로 평가하며, 슬롯 시작 시 prior uncertainty는 $e_{\text{pri}}^{(t)} = \text{tr}(\mathbf{P}_p^{(t|t-1)})$ 로 나타낸다. 즉, 작은 $\mathbf{R}^{(t)}$ 는 현재 슬롯의 $e_{\text{post}}^{(t)}$ 를 낮추고, 이어서 다음 슬롯으로 전파되는 $e_{\text{pri}}^{(t+1)}$ 도 낮춘다. 이에 따라 posterior uncertainty를 허용 범위 이하로 유지하는 것은 안전 임계형 V2X-ISAC 운용에서 반드시 만족되어야 하는 safety-critical sensing constraint가 된다.

통신 측에서는 통신 대기 비트량 $q^{(t)}$ 가

$$q^{(t+1)} = \max\{q^{(t)} - C^{(t)}, 0\} + \lambda^{(t)} \quad (4)$$

로 진화한다고 가정하며, 서비스량은 $C^{(t)} = \tau_c^{(t)} \tilde{R}^{(t)}$ 이다. 여기서 $\tilde{R}^{(t)}$ 는 유효 전송률이다. 또한 신규 유입 트래픽 $\lambda^{(t)}$ 는 통신 큐의 도착 비트 모델로서 센싱 측과는 독립적인 통신 압력을 반영하도록 구성한다. 통신 성능은

$$L^{(t)} = \tau \frac{q^{(t)}}{C^{(t)} + \epsilon} \quad (5)$$

로 정의한 communication delay metric으로 평가한다.

기지국은 자원 할당 정책 $\{a^{(t)}, \forall t \in \mathcal{T}\}$ 를 선택하여 safety-critical sensing constraint를 만족시키면서 평균 경고 지연을 최소화하는 다음과 같은 문제로 정의한다.

$$\begin{aligned} \min_{\{a^{(t)}, \forall t \in \mathcal{T}\}} \quad & \mathbb{E} \left[\frac{1}{T} \sum_{t \in \mathcal{T}} L^{(t)} \right] \\ \text{s.t.} \quad & 0 \leq a^{(t)} \leq 1, \forall t \in \mathcal{T}, \\ & e_{\text{post}}^{(t)} \leq e_{\text{th}}, \forall t \in \mathcal{T}. \end{aligned} \quad (6)$$

여기서 e_{th} 는 허용 가능한 센싱 불확실성 임계값이자 반드시 만족되어야 하는 센싱 제약의 기준이다.

iii Proposed DDPG Method

제안 기법에서는 RSU를 하나의 에이전트로 두고 연속 행동공간에서 센싱 비율 $a^{(t)}$ 를 결정한다. 여기서 행동은 슬롯 t 의 센싱 비율로서

$$a^{(t)} \in [0, 1] \quad (7)$$

로 정의되며, $1 - a^{(t)}$ 는 통신 비율을 의미한다. 상태는 센싱 측 정보와 통신 측 정보를 분리하여

$$\mathbf{s}^{(t)} = [\boldsymbol{\psi}_{\text{sen}}^{(t)}, \boldsymbol{\psi}_{\text{com}}^{(t)}]^\top \quad (8)$$

로 정의한다. 여기서 $\boldsymbol{\psi}_{\text{sen}}^{(t)}$ 와 $\boldsymbol{\psi}_{\text{com}}^{(t)}$ 는 각각 센싱 관련 관측가능 정보와 통신 관련 관측가능 정보를 의미한다. 보상은

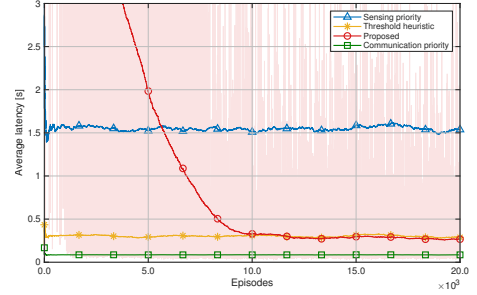
$$r^{(t)} = -\log(1 + L^{(t)}) - \lambda_{\text{exc}} \chi^{(t)} - \lambda_{\text{out}} o^{(t)} \quad (9)$$

로 정의한다. 여기서 $\chi^{(t)} = \max(0, e_{\text{post}}^{(t)} - e_{\text{th}})/e_{\text{th}}$ 는 센싱 제약의 정규화된 위반 정도를 나타내고, $o^{(t)} \in \{0, 1\}$ 는 $e_{\text{post}}^{(t)} > e_{\text{th}}$ 일 때 1이 되는 제약 위반 여부 지시변수이다. 첫 번째 항은 지연 감소를 유도하고, 나머지 두 항은 안전 임계 센싱 제약의 위반 정도와 제약 위반 발생 자체를 동시에 강하게 억제한다.

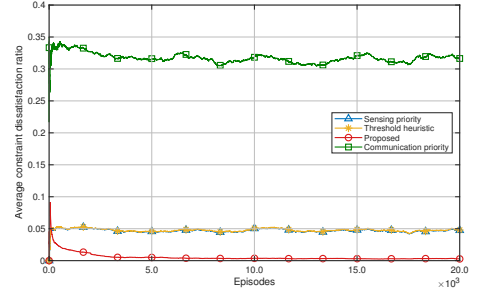
정책 학습에는 DDPG를 사용한다 [1]. Actor는 현재 상태에서부터 연속 센싱 비율을 직접 출력하고, critic은 해당 행동의 장기 성능을 평가한다.

iv Simulation Results

시뮬레이션은 교차로 환경에서 RSU가 사각지대의 단일 은닉 객체를 추적하는 동시에 주변 차량에 경고 메시지를 전송하는 안전 임계형 V2X-ISAC 상황을 가정하여 수행한다. 각 에피소드는 은닉 객체의 초기 위치와 속도, 통신 큐에 쌓인 초기 대기 비트량, 그리고 RSU와 주변 차량 사이의 경고 메시지 전송 링크 상태를 무작위로 초기화하며, 비교 기법으로는 sensing priority, communication priority, threshold heuristic, proposed DRL을 고려한다.



(a) Average episodic reward



(b) Constraint-satisfaction ratio

Figure 2: 제안 DDPG 정책의 학습 수렴 특성

Fig. 2(a)는 평균 에피소드 보상이 증가 후 안정화됨을 보이며, actor-critic 구조가 의미 있는 연속 자원 할당 정책을 학습함을 시사한다. Fig. 2(b)는 sensing satisfaction ratio가 거의 100%에 수렴함을 보여준다. 이는 proposed DRL이 safety-critical sensing constraint를 안정적으로 만족시키면서도 통신 지연을 과도하게 희생하지 않는 constraint-aware resource allocation 정책을 학습했음을 의미한다. 또한 communication priority처럼 제약을 위반하지도 않고, sensing priority처럼 필요 이상으로 보수적인 동작에 머무르지도 않는다는 점에서 의의가 있다.

ACKNOWLEDGMENT

이 연구는 2026년도 산업통상부와 한국산업기술기회평가원(KEIT)의 연구비 지원에 의한 연구임(RS-2024-KT231804)

References

- [1] T. P. Lillicrap *et al.*, “Continuous control with deep reinforcement learning,” in *Proc. ICLR*, 2016.
- [2] Y. Bar-Shalom and X. R. Li, “Maneuvering target tracking using extended Kalman filter,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, no. 3, pp. 246–254, May 1989.
- [3] F. Liu *et al.*, “Radar-assisted predictive beamforming for vehicular links: Communication served by sensing,” *IEEE Trans. Wirel. Commun.*, vol. 19, no. 11, pp. 7704–7719, Nov. 2020.
- [4] X. Meng *et al.*, “Vehicular connectivity on complex trajectories: Roadway-geometry aware ISAC beam-tracking,” *IEEE Trans. Wirel. Commun.*, vol. 22, no. 11, pp. 7408–7423, Nov. 2023.