Deep Learning Based Sensing Decision for Wireless Sensor Networks

Hewon Cho and Jungsook Bae †

† Electronics and Telecommunications Research Institute (ETRI), Daejeon, South Korea
Email: {hwcho, jsbae}@etri.re.kr

***Abstract*—In this paper, we present a sensing region selection framework for vehicular sensor networks (VSNs) for minimizing age of information (AoI) violations and transmission cost. We formulate the sensing decision problem as a Markov Decision Process (MDP), where the system state captures vehicle distribution, AoI status, and prior transmission success ratios. A reinforcement learning algorithm, Proximal Policy Optimization (PPO), is employed to learn an effective sensing decision policy that balances data freshness and communication efficiency in dynamic environments.**

I. Introduction

The evolution of 6G communications is expected to support highly intelligent, context-aware services by enabling large-scale, real-time environmental understanding [1]. As part of this vision, vehicular sensor networks (VSNs) are gaining significant attention. By equipping moving vehicles with on-board sensors, VSNs enable distributed and infrastructure-free sensing [2].

Despite these advantages, realizing effective VSNs presents two major challenges: guaranteeing the freshness of sensed data and ensuring communication efficiency under limited wireless resources. In dynamic environments, delayed or outdated information can lead to inaccurate perception of the physical world. The age of informations (AoIs) metric has been widely adopted to capture the timeliness of data [3]. Furthermore, wireless channels are subject to random fading, interference, and retransmission, which can degrade both timeliness and energy efficiency.

To address these issues, this paper proposes a deep learning based sensing decision framework for wireless VSN, formulated as Markov Decision Process (MDP). The system uses Proximal Policy Optimization (PPO) to learn optimal sensing decisions that minimize the AoI violations while reducing the transmission cost. Finally, the simulation results demonstrates the performance of the proposed sensing scheme.

II. System Model

We consider a VSN where the sensing area is partitioned into N_{S} distinct sensing regions indexed by $\mathcal{N} = \{1, \cdots, N_{\text{S}}\}$. Vehicles are distributed according to a Poisson Point Process (PPP) and can collect environmental data (e.g., temperature, humidity) within their respective regions. A single central server periodically performs sensing region decision at every sensing period denoted as T_{S}. Specifically, at each sensing round j, the central server determines the sensing region selection $a_n(j) \in \{0, 1\}$ for each sensing region $n \in \mathcal{N}$.

Let $\mathcal{V}_n(j)$ denote the set of vehicles located in sensing region n during the j-th sensing round, and $V_n(j) = |\mathcal{V}_n(j)|$ be the corresponding number of vehicles. Among these, v_n^* is defined as the vehicle closest to the geometric center of sensing region n, and is assumed to play a central role in the sensing and data reporting process. Then, the uplink Signal-to-Interference-plus-Noise Ratio (SINR) received by the roadside unit (RSU) at $\mathbf{y}_n$ from the vehicle v_n^* at $\mathbf{x}_n^*$ in sensing region n is given by

$$\gamma_{\mathbf{x}_n^*, \mathbf{y}_n} = \frac{P_{\text{tx}} h_{\mathbf{x}_n^*, \mathbf{y}_n} \ell_{\mathbf{x}_n^*, \mathbf{y}_n}}{\sigma^2 + I_n}, \tag{1}$$

where P_{tx} denotes the transmission power of the vehicle, $h_{\mathbf{x}_n, \mathbf{y}_n}$ and $\ell_{\mathbf{x}_n, \mathbf{y}_n}$ are the channel fading gain and the pathloss between the vehicle and the RSU, respectively, σ^2 is the Additive White Gaussian Noise (AWGN) power, and I_n is the inter-cell interference from users using same uplink frequency band with vehicle v_n^*. Then the successful uplink transmission for sensing region n at time t is given by

$$U_n(t) = \begin{cases} 1, & \text{if } B \log_2\left(1 + \gamma_{\mathbf{x}_n^*, \mathbf{y}_n}\right) > \delta, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where B is the channel bandwidth and δ is the target data rate. Here, $\delta = \frac{D}{T_{\text{tx}}}$ where D is the sensed data size and T_{tx} is the transmission time.

To ensure reliable transmission, the system permits up to N_{rtx} retransmissions per sensing round in the event of a transmission failure. Under the retransmission scheme, the success of uplink retransmission for sensing region n at j-th sensing round can be modeled as

$$R_n(j) = \begin{cases} 1, & \text{if } \sum_{c=jT_{\text{S}}}^{jT_{\text{S}} + N_{\text{rtx}} T_{\text{tx}}} U_n(c) = 1, \\ 0, & \text{otherwise}. \end{cases} \tag{3}$$

Additionaly, the total number of retransmission attemps required in the j-th sensing round, denoted as $r_n^o(j)$, can be expressed as $r_n^o(j) = \min\left\{k \left| U_n\left(jT_{\text{S}} + (k-1)T_{\text{tx}}\right) = 1, \forall k \in \{1, \cdots, N_{\text{rtx}}\}\right.\right\}$.

Therefore, the overall success of the data update process is formulated as

$$S_n(j) = \begin{cases} 1, & \text{if } a_n(j)\, V_n(j)\, R_n(j) > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

At time t, the AoI at the RSU for sensing region n is defined as $\Delta_n(t) = t - t_{\mathrm{g},n}$ where $t_{\mathrm{g},n}$ denotes the generation time of the most recently updated data of the sensing region n. Once a sensing decision is made by the central server, there exists a fixed decision latency denoted as T_{d}, which contributes to the increase of the AoI.

When the sensed environmental data follows a stationary Gaussian process, the correlation coefficient between the current value and the outdated sensed data from sensing region n is given by

$$\rho_n(\Delta_n(t)) = \exp(\beta_n \Delta_n(t)), \tag{5}$$

where β_n is the weight of the temporal error for sensing region n [4]. Then, the estimation error is obtained as

$$\epsilon_n(\Delta_n(t)) = 1 - (\rho_n(\Delta_n(t)))^2 = 1 - \exp(-2\beta_n \Delta_n(t)). \tag{6}$$

The estimated data for sensing region n is considered invalid if the error exceeds a predefined threshold θ_n, i.e., $\epsilon_n(\Delta_n(t)) \geq \theta_n$. Then, we obtain the AoI violation condition as follows.

$$\Delta_n(t) \geq A_{\mathrm{th},n} = -\frac{\ln(1-\theta_n)}{2\beta_n}. \tag{7}$$

Finally, the error violation in sensing region n at j-th sensing round, denoted as $g_{n,j}$, is defined as the duration during which the AoI exceeds the threshold $A_{\mathrm{th},n}$.

III. MDP Formulation

The sensing region selection process at each sensing round is formulated as a MDP to jointly optimize information freshness and communication efficiency. The system **state** at the j-th sensing round is defined as

$$\mathbf{s}(j) = \{\mathbf{T}(j), \mathbf{V}(j-1), \mathbf{A}^*(j)\} \tag{8}$$

where $\mathbf{T}(j) = [T_1(j), \cdots, T_{N_{\mathrm{S}}}(j)]$ denotes the successful transmission ratio at sensing round j and is computed as

$$\mathbf{T}(j) = \begin{cases} -1, & \text{if } \sum_{i \in \mathcal{J}_{\mathrm{w}}(j)} a_n(j) = 0, \\ \frac{\sum_{i \in \mathcal{J}_{\mathrm{w}}(j)} U_n(j)}{\sum_{i \in \mathcal{J}_{\mathrm{w}}(j)} r_n^o(j)}, & \text{otherwise.} \end{cases} \tag{9}$$

where $\mathcal{J}_{\mathrm{w}}(j) = \{i \mid \max(1, j-W) \geq i \geq j-1\}$ where W is the window size. In the system state, $\mathbf{V}(j-1) = [V_1(j-1), \cdots, V_{N_{\mathrm{S}}}(j-1)]$ denotes the number of vehicles in previous round and $\mathbf{A}^*(j) = [A_1(j), \cdots, A_{N_{\mathrm{S}}}(j)]$ denotes the current AoI violation status where $A_n(j) = A_{\mathrm{th},n} - \Delta_n(j)$.

The **action** at round j, denotes as

$$\mathbf{a}(j) = \{a_1(j), \cdots, a_{N_{\mathrm{S}}}(j)\}, \tag{10}$$

corresponds to the sensing region selection.

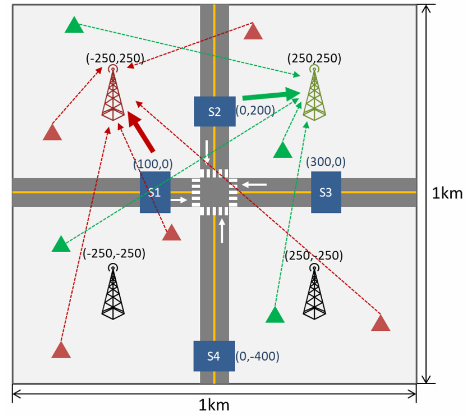

Fig. 1. Simulation model.

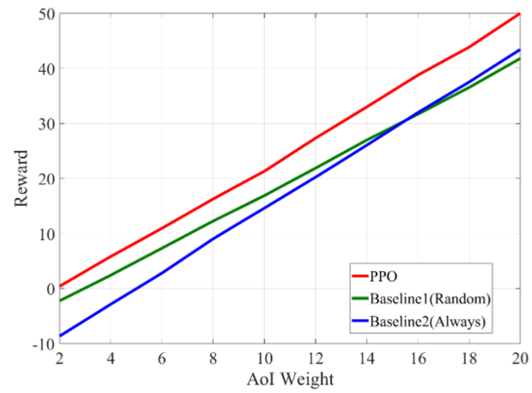

Fig. 2. Reward versus the AoI weight when $\omega_{\mathrm{tx}} = 2$.

The **reward function** is designed to balance AoI freshness and transmission cost. Hence, the immediate reward at round j is defined as

$$\mathbf{r}(j) = \omega_{\mathrm{aoi}} \sum_{n=1}^{N_{\mathrm{S}}} \frac{T_{\mathrm{S}} - g_{n,j}}{N_{\mathrm{S}}} - \omega_{\mathrm{tx}} \sum_{n=1}^{N_{\mathrm{S}}} \frac{P_{\mathrm{tx}} T_{\mathrm{tx}} S_n(j)\, r_n^o(j)}{N_{\mathrm{S}}}, \tag{11}$$

where ω_{aoi} and ω_{tx} are weighting coefficients for AoI violation and transmission power, respectively.

IV. Numerical Results

For performance evaluation, we consider a simulation environment as illustrated in Fig.1, consisting of 4 RSUs and 4 sensing regions. Vehicles are spatially distributed according to a PPP and the interfering user density is set to 10^{-7} nodes/m^2. We choose $D = 2.5$KB, $P_{\mathrm{tx}} = 20$dBm, $\sigma = -100$dBm, $B = 5$MHz, $\delta = 2$Mbps, $T_{\mathrm{tx}} = 10$ms, $T_{\mathrm{S}} = 1$ms, $T_{\mathrm{d}} = 0.1$ms, $N_{\mathrm{rtx}} = 3$, and $\theta_n = 0.03$.

Two baseline schemes are used for comparison: (i) random sensing, where each sensing region is selected with a probability of 0.5 at each sensing round; and (ii) always sensing, where all sensing regions are selected in every round regardless of context.

Figure 2 shows the variation in the reward as a function of ω_{aoi}, under a fixed transmission cost weight of $\omega_{\mathrm{tx}} = 2$. The proposed scheme, implemented using the PPO algorithm–a deep reinforcement learning method–achieves significantly higher reward compared to both baseline schemes, demonstrating its effectiveness in balancing AoI minimization and transmission efficiency.

V. Conclusion

This paper presents a reinforcement learning-based sensing decision for vehicular sensor networks aimed at minimizing AoI violations and reducing transmission cost. By modeling the problem as a MDP and leveraging the PPO algorithm, the proposed framework adaptively selects sensing regions based on system dynamics such as vehicle number, AoI status, and transmission success rates. Finally, through numberical results, we show that the proposed scheme achieves better performance compared to the baseline schemes.

VI. Acknowledgment

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-00972, Development of Intelligent Wireless Access Technologies).

References

[1] H. Viswanathan and P. E. Mogensen, "Communications in the 6g era," *IEEE Access*, vol. 8, pp. 57 063–57 074, 2020.

[2] Q. Yuan, H. Zhou, Z. Liu, J. Li, F. Yang, and X. Shen, "Cesense: Cost-effective urban environment sensing in vehicular sensor networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 9, pp. 3235–3246, 2018.

[3] S. Yun, D. Kim, C. Park, and J. Lee, "Learning-based sensing and computing decision for data freshness in edge computing-enabled networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11 386–11 400, 2024.

[4] M. L. Stein, "Space–time covariance functions," *J. Amer. Stat. Assoc.*, vol. 100, no. 469, pp. 310–321, 2005.