Disturbance-Resilient Agentic AI Control for Age-Fair SPS in 5G V2I Networks

Junchen Bai, Qiong Wu, *Senior Member, IEEE*, Kezhi Wang, *Senior Member, IEEE*, Pingyi Fan, *Senior Member, IEEE*, Guoqiang Mao, *Fellow, IEEE*, and Khaled B. Letaief, *Fellow, IEEE*

***Abstract*—In fifth-generation (5G) New Radio (NR) vehicle-to-infrastructure (V2I) networks, semi-persistent scheduling (SPS) reduces signaling but may retain an unsuitable resource reservation interval (RRI) after the local load changes. We study age-fair keep/reselect control and distinguish a reward-aligned age-fair utility (RAAFU) surrogate from direct age-of-information (AoI), Jain freshness, and SPS radio-resource collision endpoints. The Full agent combines deterministic Rule/model-predictive-control (MPC) evaluation and a safety gate with an event-triggered large-language-model (LLM) proposal path. Across 128 independent scenario families, Full achieved the highest displayed Jain freshness, 0.9386, and a 1.009% collision rate; it also slightly reduced mean/q95 AoI relative to the same-stack controls. In an independent 256-case shock study, the LLM path reduced q95 AoI and collision relative to a budget-matched non-LLM selector, showing a measurable component contribution, although the stronger Rule remained competitive. The main trade-off is a higher absolute AoI than some external comparators, and measured provider latency is on the order of seconds rather than slots. These findings support only a simulation-based slow-path role for the LLM; they do not establish real-time deployment.**

***Index Terms*—5G NR V2I, semi-persistent scheduling, age fairness, agentic AI, disturbance adaptation.**

I. Introduction

V2I communication supports connected driving through periodic state updates [1], [2]. AoI is the time elapsed since the generation of the freshest delivered update, PAoI is sampled immediately before a successful update, and age fairness describes how evenly freshness is distributed across users [3], [4]. Because throughput or packet delay alone does not measure these properties, we report direct mean/q95 AoI and Jain freshness and use RAAFU as an engineering control utility [5], [8], [9].

At the MAC layer, SPS saves signaling by reusing a reservation, but persistence may retain an unsuitable RRI after traffic density, collision risk, or surrounding reselection changes [10]. Recent NR vehicular work addresses aperiodic traffic, age-aware congestion, and the freshness effects of reservation reuse [11], [12], [13]. The remaining question is how to adapt RRI while preserving both age fairness and the signaling advantage of SPS.

Vehicular resource allocation now includes learning-based, graph-based, and age-aware control [14], [15], [16], [17]. Recent studies also combine velocity-adaptive access with joint fairness and AoI optimization [5], [6], while multi-UAV-assisted IoV work jointly considers trajectory control, resource allocation, and task offloading [7]. Under disturbed SPS operation, however, one compressed causal observation may represent several hidden reservation states. The controller must compare candidate RRIs while preserving a deterministic fallback.

We study an *agentic artificial intelligence (agentic AI)* implementation for age-fair SPS control. The Full agent combines deterministic SPS tools, a risk gate, a fallback, and an event-triggered LLM proposal path. The ablation study shows that this LLM path contributes to candidate ordering under synthetic shocks; it is neither the slot-level executor nor the final safety authority.

Our contribution is threefold.[1] First, we clarify the constrained keep/reselect formulation and distinguish its RAAFU surrogate from direct AoI and fairness endpoints. Second, we implement a two-time-scale Full agent: Rule/MPC provides the fast path, while an event-triggered LLM provides bounded candidate ordering under a fixed budget. Third, we evaluate direct endpoints, same-stack controls, weight sensitivity, and end-to-end latency under matched disturbance conditions. These comparisons show the Full gains while stating the measured trade-offs and limits.

II. System Description and Modeling

A. System Model

We consider a single-cell 5G NR V2I link using SPS at the MAC layer. Only vehicle 0 runs the agentic controller. It chooses a keep/reselect branch and a candidate $r_t \in \mathcal{R}$; keep simply reuses the previous reservation. Other vehicles act as background SPS users and shape the local load, collision risk, and hidden reservation dynamics.

Received 28 August 2026; accepted 4 September 2026. Date of publication 9 September 2026; date of current version 21 September 2026. This work was supported in part by the National Natural Science Foundation of China under Grant Nos. 62671277, 62531015, and U25A20399, in part by the Basic Research Program of Jiangsu under Grant No. BK20252084, in part by the National Key Research and Development Program of China under Grant No. 2021YFA1000500(4), in part by the Shanghai Kewei under Grant No. 24DP1500500, and in part by the Hong Kong Research Grant Council under the Areas of Excellence (AoE) Scheme under Grant No. AoE/E-601/22-R. The associate editor coordinating the review of this letter and approving it for publication was A. Moubayed. *(Corresponding author: Qiong Wu.)*

Junchen Bai and Qiong Wu are with the School of Automation and Intelligent Science (School of IoT), Jiangnan University, Wuxi 214122, China (e-mail: junchenbai@stu.jiangnan.edu.cn; qiongwu@jiangnan.edu.cn).

Kezhi Wang is with the Department of Computer Science, Brunel University London, UB8 3PH Uxbridge, U.K. (e-mail: Kezhi.Wang@brunel.ac.uk).

Pingyi Fan is with the Department of Electronic Engineering, State Key Laboratory of Space Network and Communications, and Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China (e-mail: fpy@tsinghua.edu.cn).

Guoqiang Mao is with the School of Transportation, Southeast University, Nanjing 210096, China (e-mail: g.mao@ieee.org).

Khaled B. Letaief is with the Department of Electrical and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong (e-mail: eekhaled@ust.hk).

This article has supplementary downloadable material available at https://doi.org/10.1109/LCOMM.2026.3732043, provided by the authors.

Digital Object Identifier 10.1109/LCOMM.2026.3732043

[1]Source code: github.com/qiongwu86/Disturbance-Resilient-Agentic-AI-Control-for-Age-Fair-SPS-in-5G-V2I-Networks

1558-2558 © 2026 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Southeast University. Downloaded on October 30,2026 at 12:09:54 UTC from IEEE Xplore. Restrictions apply.

Let V_t be the active set and $N_v(t) = |V_t|$ in a road segment of length L, with density $\varrho_t = N_v(t)/L$. Vehicle count follows

$$\begin{aligned} &\Pr\{\Delta N_t = 1\} = p_a, \quad \Pr\{\Delta N_t = -1\} = p_d, \\ &\Pr\{\Delta N_t = 0\} = 1 - p_a - p_d, \\ &N_v(t+1) = \left[N_v(t) + \Delta N_t\right]_{N_{\min}}^{N_{\max}}, \end{aligned} \tag{1}$$

where λ_v and μ_v are the arrival and departure rates in s^{-1}, $T_{\mathrm{eval}} = 1$ s is the evaluation interval, $c = \max\{1, (\lambda_v + \mu_v)T_{\mathrm{eval}}\}$, $p_a = \lambda_v T_{\mathrm{eval}}/c$, and $p_d = \mu_v T_{\mathrm{eval}}/c$. Thus, $p_a, p_d \in [0,1]$ and $p_a + p_d \le 1$. The clipping operator is applied after the sampled event and enforces $N_{\min} \le N_v(t+1) \le N_{\max}$. Within an evaluation interval, the committed controlled RRI is applied first, the background RRIs transition, the count event is realized, and the state/age summaries are then updated. A boundary arrival or departure that would violate the range is absorbed by clipping.

For RRIs, $r_0(t) = r_t$, while each background vehicle $i \in V_t \setminus \{0\}$ has its own $r_i(t)$. Static cases keep background RRIs fixed. Disturbed cases use a Markov-driven process over $\mathcal{S}_R = \{s_1, \ldots, s_K\}$, motivated by SPS reservation analysis [18]:

$$r_i(t) = s_{m_i(t)}, \quad m_i(t+1) = \Pi_{[1,K]}(m_i(t) + d_i(t)\xi_i(t)), \tag{2}$$

where $d_i(t) \in \{-1, 1\}$ is a persistent ping-pong direction, $\xi_i(t) \sim \mathrm{Bernoulli}(p_s)$, and p_s is a transition probability per evaluation step. The operator $\Pi_{[1,K]}$ reflects an attempted transition: at K it returns $K-1$ and sets the next direction to -1, while at 1 it returns 2 and sets the next direction to + 1. A simulation shock at τ adds ΔN background vehicles and may reset their RRI mode, Markov states, and p_s; only r_t is controlled. Because vehicle $0 \in V_t$, the active set is nonempty, and the effective RRI in milliseconds is $\bar{r}_t = \left(N_v(t)^{-1} \sum_{i \in V_t} r_i(t)^{-1}\right)^{-1}$.

The state s_t summarizes density, RRI context, collision indicators, and recent age information. Since the segment is treated as one local sensing neighborhood, $N_{\mathrm{comp}}(t) = N_v(t)$. Following SPS collision modeling [18],

$$\begin{aligned} &N_{\mathrm{res}}(t) = \frac{\bar{r}_t n_{\mathrm{RBG}}}{T_{\mathrm{slot}}}, \quad q_t = \left[\frac{\alpha_{\mathrm{sps}}}{N_{\mathrm{res}}(t)}\right]_0^1, \\ &p_{\mathrm{coll}}(t) = 1 - (1 - q_t)^{N_{\mathrm{comp}}(t)-1}. \end{aligned} \tag{3}$$

Here, n_{RBG} is the number of resource block groups per RRI cycle, T_{slot} and $\bar{r}_t$ use the same time unit, $N_{\mathrm{res}}(t) > 0$, and $\alpha_{\mathrm{sps}} \ge 0$ is a resource-occupancy factor. We set $N_{\mathrm{comp}}(t) = N_v(t)$ as a single-neighborhood modeling assumption. Clipping makes $p_{\mathrm{coll}} \in [0,1]$ and prevents a negative binomial base. This is a simulated SPS radio-resource collision probability, not a vehicle-crash probability.

We model peak age of information (PAoI) through SPS waiting and retransmission delays, following age of information (AoI) and persistence-aware NR-V2X analysis [8], [13]. The two-state SPS abstraction is

$$\begin{aligned} &\eta_t = \frac{p_0}{q-p+1} + \left(1 - \frac{1}{q-p+1}\right)(1 - p_{\mathrm{coll}}(t)), \\ &P_{\mathrm{fail}}(t) = p_{\mathrm{coll}}(t)(1 - \eta_{\mathrm{re}}), \quad \pi_{1,t} = \frac{1-\eta_t}{2 - \eta_t - P_{\mathrm{fail}}(t)}. \end{aligned} \tag{4}$$

Here, $0 \le p_0, \eta_{\mathrm{re}} \le 1$, $1 \le p \le q$ are integers, and $0 \le p_{\mathrm{coll}} \le 1$. The two states represent retained and re-evaluated reservation conditions; η_t is the retained-state transition term, P_{fail} is the SPS failure term, and $\pi_{1,t}$ is the stationary probability of state 1. The ratio is defined when $2 - \eta_t - P_{\mathrm{fail}}(t) > 0$; this condition holds for $p_0 = 0.8$, $(p,q) = (5,15)$, and $\eta_{\mathrm{re}} = 0.7$ [13], [18]. The analytic delay and age surrogate are

$$\begin{aligned} &D_t(t) = T_s + \frac{T_c\, p_{\mathrm{coll}}(t)}{1 - p_{\mathrm{coll}}(t)}, \\ &A_t = 1.5\bar{r}_t + \pi_{1,t}\left(T_{\mathrm{sensing}} + \frac{wT_{\mathrm{slot}}}{2}\right) + D_t(t), \end{aligned} \tag{5}$$

where all time quantities are in milliseconds. The $1.5\bar{r}_t$ term approximates one RRI of waiting plus a half-RRI phase offset. For $p_{\mathrm{coll}} = 1$, the implementation uses the bounded fallback $D_t = T_s + 3T_c$ rather than evaluating the singular ratio. We denote this analytic score input by A_t^{sur} because it is not the direct realized AoI endpoint.

For direct evaluation, let $g_i(u)$ be the generation time of the freshest update from user i delivered by continuous time u. Its current AoI is $\Delta_i(u) = u - g_i(u)$. A success first records the pre-reset value as a PAoI event and then sets g_i to that packet's generation time; a failure leaves g_i unchanged. Transmission outcomes are processed before a departure at the same time, followed by arrival, packet generation, measurement delivery, control opportunity, and commitment. Over the analysis window $\mathcal{W}$ and eligible set $V(u)$, we report

$$\begin{aligned} \overline{\Delta} &= \frac{1}{|\mathcal{W}|}\int_{\mathcal{W}} \frac{1}{|V(u)|} \sum_{i \in V(u)} \Delta_i(u)\, du, \\ \Delta_{.95} &= \frac{1}{|\mathcal{W}|}\int_{\mathcal{W}} Q_{.95}\{\Delta_i(u) : i \in V(u)\}\, du, \\ J_f &= \frac{\left(\sum_{i \in V_{\mathcal{W}}} f_i^{\mathrm{dir}}\right)^2}{|V_{\mathcal{W}}| \sum_{i \in V_{\mathcal{W}}} (f_i^{\mathrm{dir}})^2}, \quad f_i^{\mathrm{dir}} = \frac{1\ \mathrm{ms}}{1\ \mathrm{ms} + \overline{\Delta}_i}, \\ R_{\mathrm{coll}} &= N_{\mathrm{coll}}/N_{\mathrm{tx}}. \end{aligned} \tag{6}$$

Here, $\overline{\Delta}_i$ is the user's time-average AoI over its eligible presence, q95 is the cross-user age quantile integrated over time, and event PAoI is separately averaged over users with at least one success. These direct endpoints do not equal RAAFU.

B. Objective Function

Using the SPS and PAoI quantities above, we define the reward-aligned age-fair utility (RAAFU). RAAFU is the utility optimized here, not a formula taken directly from prior work. For vehicle i, we define freshness as $f_i(t) = 1/(\Delta_i(t) + \epsilon)$. The age-fair factor, RRI-efficiency factor, and collision factor are

$$\begin{aligned} \Gamma_i(t) &= \exp\left(-k_{\mathrm{age}}\left|\log\frac{f_i(t)}{\bar{f}(t)}\right|\right), \\ E_{\mathrm{rri}}(r,t) &= \exp\left(-\alpha_{\mathrm{rri}}\left|\log\frac{r}{r_{\mathrm{opt}}(N_v(t))}\right|\right), \end{aligned} \tag{7}$$

$$C(t) = \exp(-\beta_{\mathrm{coll}}\, p_{\mathrm{coll}}(t)), \tag{8}$$

Authorized licensed use limited to: Southeast University. Downloaded on October 30,2026 at 12:09:54 UTC from IEEE Xplore. Restrictions apply.

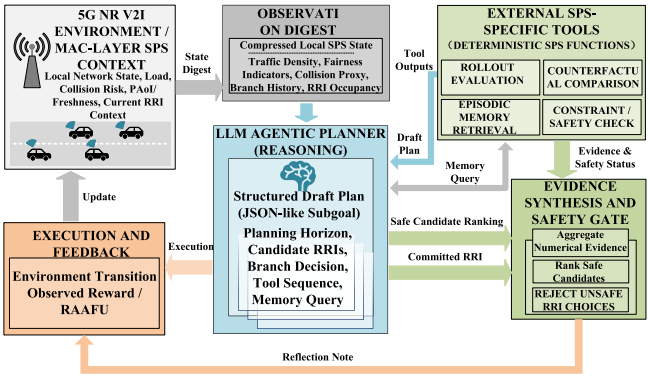


Fig. 1. Agentic SPS control architecture.

The step utility is

$$U_t(r_t) = \frac{1}{|V_t|} \sum_{i \in V_t} \Gamma_i(t)\, E_{\mathrm{rri}}(r_t, t)\, C(t), \tag{9}$$

where $\bar{f}(t) = |V_t|^{-1} \sum_{i \in V_t} f_i(t)$, $\epsilon = 10^{-6}$ in the implemented age unit, $k_{\mathrm{age}} = 0.6$, $\alpha_{\mathrm{rri}} = 0.1$, and $\beta_{\mathrm{coll}} = 0.2$ are the operating parameters; $r_{\mathrm{opt}}(N_v)$ is an offline density mapping. The collision term is evaluated after substituting the candidate $r_0(t) = r_t$ into p_{coll}. With a nonempty active set, all three factors and U_t are strictly positive and at most one. RAAFU favors balanced inverse age, an RRI near the calibrated reference, and low modeled collision risk, but it is a composite surrogate rather than a direct AoI or fairness estimand.

C. Optimization Problem

At each step, the controller selects the branch and RRI of the controllable node to maximize cumulative utility while preserving SPS persistence when keep is selected. The optimization problem is

$$\begin{aligned} \max_{\pi} \quad & \mathbb{E}_{\pi,\mathcal{M}}\left[\sum_{\tau=t}^{t+H-1} U_\tau(r_\tau) \mid \mathcal{H}_t\right] \\ \text{s.t.} \quad & (b_t, r_t) = \pi(\mathcal{H}_t), \quad b_t \in \{\text{keep}, \text{reselect}\}, \\ & r_t \in \mathcal{R}, \quad r_t = r_{t-1} \text{ if } b_t = \text{keep}, \\ & \widehat{p}_{\mathrm{coll}}(r_t \mid \mathcal{H}_t, \mathcal{M}) \le p_{\max}, \end{aligned} \tag{10}$$

where H is the planning horizon, $\mathcal{H}_t$ contains only causally delivered measurements and branch history, $\mathcal{M}$ is the controller model, and $p_{\max}$ is the configured predicted-risk threshold. We set $p_{\max} = 0.25$ for the direct-endpoint evaluation and 0.35 for the controlled ordering ablation; the two studies are analyzed separately. If no candidate is feasible, the deterministic fallback uses a safe backup or selects the largest legal RRI. The realized collision endpoint is reported separately from this predicted feasibility constraint.

III. Proposed Agentic SPS Controller

Section II leaves the controller with a causal but compressed SPS view. We therefore use the Full controller shown in Fig. 1. We used OpenAI GPT-5.4-mini for the Full and Plain-LLM paths. The observation digest enters the bounded LLM planner, after which the SPS-specific tools evaluate legal candidates. Evidence synthesis and the safety gate commit the RRI. Thus, the LLM participates in planning, while Rule/MPC remains the final gate and fallback.

At each SPS decision step, the controller forms

$$\begin{aligned} o_t &= \{N_v(t), r_t^{\mathrm{cur}}, A_t^{\mathrm{sur}}, P_t, F_t, U_t, d_t, b_{t-1}, \rho_{t-1}, \zeta_{t-1}\}, \\ p_t &= \{g_t, H_t, b_t, \mathcal{C}_t, \mathcal{T}_t, q_t^{\mathrm{mem}}, \phi_t, \kappa_t\}, \end{aligned} \tag{11}$$

Here, $N_v(t)$ is the delivered active count, r_t^{cur} is the committed RRI, and A_t^{sur}, P_t, F_t, and U_t summarize the analytic age input, modeled collision, fairness, and RAAFU. The remaining fields record branch history, reflection hints, the bounded plan, fallback, and non-operational confidence. The detector d_t uses only delivered count/ledger, channel-busy, decode, collision, or age-residual summaries; the controller receives no shock label. Detector false alarms, misses, and delay require further evaluation. The first pass has at most four legal RRIs, and one replan may expand it to five. Invalid or ineligible proposals leave Rule/MPC in control.

Five deterministic tools support the decision: state_digest, memory_query, rollout_candidates, counterfactual, and constraint_check. Memory retrieves at most three observable-bucket records. Rollout returns $\{\hat{A}_t, \hat{P}_t, \hat{F}_t, \hat{U}_t, G_t\}$, where $\hat{A}_t(r)$ denotes the predicted worst-user age for candidate r. $\{\hat{P}_t, \hat{F}_t, \hat{U}_t, G_t\}$, and the constraint tool forms $\mathcal{C}_t^{\mathrm{safe}} = \{r : \hat{P}_t(r) \le p_{\max}\}$. Rollout always runs, whereas memory is conditional; one score-guided replan may expand the set to five. The Full agent includes the LLM ordering path because the controlled ablation shows a measurable gain over the budget-matched selector. Since the rollout model and simulator use the same nominal SPS abstraction, the reported results do not characterize controller-model mismatch.

After the tools return, the controller ranks the legal RRIs by combining rollout, memory, and counterfactual evidence. For each candidate r, it computes

$$\begin{aligned} S_t^{\mathrm{roll}}(r) &= 0.50\,\hat{U}_t(r) + 0.20\,\hat{F}_t(r) + 0.15 e^{-\hat{A}_t(r)/1200} \\ &\quad + 0.10\left(1 - \hat{P}_t(r)\right) + 0.05\left(\hat{U}_t(r) - U_t\right), \\ S_t^{\mathrm{synth}}(r) &= S_t^{\mathrm{roll}}(r) + B_t^{\mathrm{mem}}(r) + B_t^{\mathrm{cf}}(r), \\ r_t^{\star} &= \arg\max_{r \in \mathcal{C}_t^{\mathrm{safe}}} S_t^{\mathrm{synth}}(r), \end{aligned} \tag{12}$$

The five weights in Eq. (12) define the operating point $W_0 = (0.50, 0.20, 0.15, 0.10, 0.05)$. They were selected in a preliminary sweep before evaluation. A 160/160 calibration–holdout sensitivity study found that all 76 local sum-preserving perturbations kept the same action, but the minimum stability across 4,195 global configurations fell to 64.375%. Thus, W_0 is locally stable but globally sensitive. The implemented memory score uses $\mathrm{RAAFU} + 0.35F - 8P$ before clipping to $[-0.08, 0.08]$. Its $-8P$ term reached the lower clip for all 800 evaluated candidates and did not affect their ranking. In this cohort, collision control therefore came from the explicit risk gate rather than this clipped term.

The final deterministic commit emits the chosen/backup RRI, branch, expected gain/risk, decision records, and a replan flag. An empty, inconsistent, or unsafe proposal falls back to the score-ranked result or a safe high RRI. The learning reward trains the policies and updates reflection records; it is

Authorized licensed use limited to: Southeast University. Downloaded on October 30,2026 at 12:09:54 UTC from IEEE Xplore. Restrictions apply.

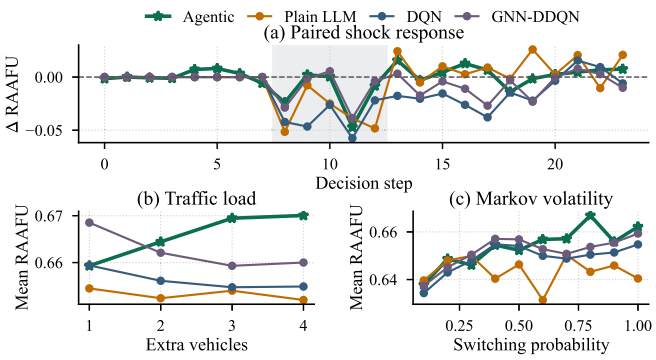


Fig. 2. RAAFU under a paired step disturbance and under traffic-load and Markov-volatility sweeps. RAAFU is a composite control utility rather than a direct AoI metric.

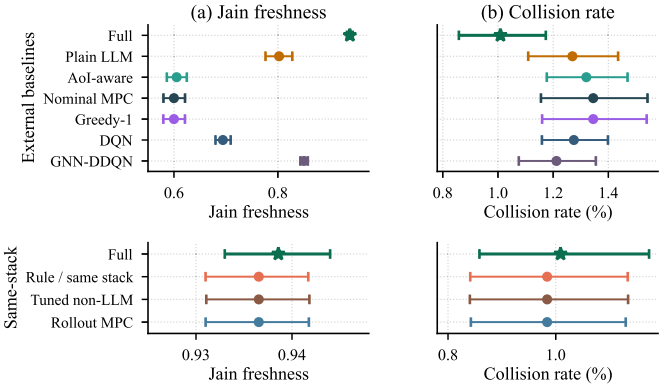


Fig. 3. Direct endpoints over 128 independent scenario families. Upper axes show external/standard baselines; lower axes magnify the four same-stack methods. Bars are descriptive 95% scenario-family CIs; collision is an SPS radio-resource event rate.

separate from the reported endpoints. Its complete expression is provided in the Supplementary Material.

All baselines use the same causal frames and legal-action generator. Rule/same stack substitutes a hand-coded ordering, TunedNonLLM is budget matched, Rollout MPC uses the same evaluator without an LLM proposer, and Greedy-1 uses one-step rollout utility. The decision-critical Plain-LLM instruction is: *"Use only the causal observation and eligible actions in the supplied JSON. You have no tools, no memory, no plant truth, no scenario or arm identity, no shock label, no future tape, and no method-ranking information. Rank every eligible action once, then choose the first-ranked action. Return exactly one JSON object matching the required output schema and no other text. Never invent an action, target RRI, observation, or future outcome."* The user instruction requires an exact permutation of all eligible actions and commits the first item. Temperature and retry budget are zero; malformed or ineligible output fails closed. In Full, the LLM instead selects one eligible high-level program transition, never an RRI; the deterministic evaluator and gate commit the action. The complete prompts, schema keys, rationale codes, hashes, and fallback contracts are provided in the Supplementary Material.

IV. Performance Evaluation

We compare Full with three same-stack controls and six external or standard baselines. The evaluation includes RAAFU disturbance responses, direct endpoints over 128 independent scenario families, and controlled LLM/MPC components. Fig. 2 first shows that Full stays closer to the matched reference over most of the paired shock window and moves ahead at medium-to-high traffic load and Markov volatility. The curves are descriptive. Since RAAFU combines age balance, RRI efficiency, and modeled collision, direct AoI, fairness, and collision are examined next.

The direct-endpoint evaluation uses 128 independent scenario families under matched disturbed/reference traces; DQN and GNN-DDQN aggregate five training seeds within each family. Table I gives the endpoint values, while Fig. 3 shows Jain freshness and collision with 10,000 scenario-family bootstrap replicates. Simultaneous intervals and Holm-adjusted comparison families are described in the Supplementary Material.

TABLE I

Direct Endpoints Under Disturbance ($n = 128$ Independent Scenario Families). Lower AoI and Collision Are Favorable, as Is Higher Jain Freshness. CIs Are Reported in Fig. 3 and Table S1

Method	mean AoI	q95 AoI	Jain	Coll. (%)
Full	372.267	856.379	0.9386	1.009
Rule/same stack	374.420	859.594	0.9365	0.984
Tuned non-LLM	374.420	859.594	0.9365	0.984
Rollout MPC	374.420	859.594	0.9365	0.984
Plain LLM	343.487	885.312	0.8023	1.270
AoI-aware	318.974	812.373	0.6051	1.320
Nominal MPC	330.918	906.447	0.5999	1.345
Greedy-1	330.918	906.447	0.5999	1.345
DQN	320.797	808.463	0.6941	1.275
GNN-DDQN	327.188	805.393	0.8503	1.212

A. Direct Freshness and Fairness

Table I shows that Full has the highest displayed Jain freshness, 0.93858 (CI $[0.93300, 0.94397]$). Its mean AoI is 372.267 ms, slightly lower than those of the three same-stack controls but higher than those of the six external/standard comparators. This is a cross-user balance–absolute-age tradeoff, not a uniform AoI improvement. As shown in Fig. 3, Full has favorable same-stack fairness and AoI point estimates. Its collision rate is 1.009% (CI $[0.858\%, 1.173\%]$), slightly above the other three same-stack point estimates, and the overlapping CIs do not establish a universal LLM advantage.

B. Controlled Components and Sensitivity

Fig. 4 separates the LLM ordering path from the deterministic MPC/evaluator. In panels (a)–(b), LLM minus TunedNonLLM is -135.286 ms for q95 AoI (CI $[-159.025, -105.275]$) and -0.061987 for the collision rate (CI $[-0.072841, -0.048010]$), where the collision-rate difference and its CI are expressed as fractions and the point estimate is equivalent to -6.1987 percentage points; 15/16 batches agree, and Holm-adjusted $p = 0.0001221$ for both endpoints. Mean AoI also falls by 81.153 ms, although fairness worsens by 0.05963. The stronger Rule matches the LLM in 247/256 cases; LLM–Rule

Authorized licensed use limited to: Southeast University. Downloaded on October 30,2026 at 12:09:54 UTC from IEEE Xplore. Restrictions apply.

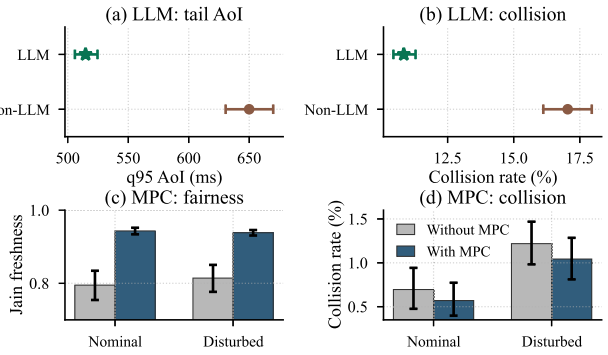


Fig. 4. Controlled components. (a)–(b) LLM versus budget-matched selector (256 cases, 16 provider batches); (c)–(d) MPC/evaluator over 64 paired scenario families per regime. Full effects and AoI costs are in Table S2.

is + 12.584 ms for q95 AoI and + 0.005833 for collision. The LLM path therefore has a measurable contribution relative to the budget-matched selector, while Rule remains the stronger deterministic comparator. These results do not attribute every Full-agent gain to the LLM.

Panels (c)–(d) show the effect of the MPC/evaluator. Jain freshness rises from 0.7950 to 0.9432 (nominal) and from 0.8142 to 0.9383 (disturbed), while collision falls from 0.6957% to 0.5700% and from 1.2187% to 1.0430%. This improvement comes with mean/q95 AoI costs of $52.607/61.385$ and $39.231/42.397$ ms. The 4,272-configuration sensitivity study in Table S3 gives 86.63% global acceptability and a minimum action stability of 64.375%, indicating local stability around W_0 but appreciable sensitivity over the global weight range.

C. Latency and Limitations

A separate profile included 64 completed ProgramProposal decisions and 64 completed Plain-LLM decisions. Total-decision p50/p95/p99 values were $3255/14036/14496$ ms and $18984/22257/26669$ ms, respectively; tool/candidate-rollout p95 values were 1142 ms for ProgramProposal and 1242 ms for Plain LLM, respectively. The measured seconds-scale delay is incompatible with a 0.5–1-ms NR slot. Within Full, the LLM therefore runs as an asynchronous slow path while Rule/MPC keeps the control loop nonblocking.

The prompts, schemas, fallback rules, and additional statistical details are provided in the Supplementary Material. Detector false alarms, misses, and delay have not been measured. The rollout model is nominal, and the global weight study shows appreciable sensitivity. Real-time deployment will require detector validation and evaluation under controller-model mismatch.

V. Conclusion

This letter casts age-fair SPS control as a constrained keep/reselect problem and separates direct AoI, fairness, and SPS collision endpoints from RAAFU. Across the evaluated conditions, Full gives the highest displayed Jain freshness and slightly lower AoI than the same-stack controls. Its integrated LLM path also improves q95 AoI and collision over the budget-matched selector in the controlled ablation. The costs are a higher absolute AoI than some external comparators and seconds-scale provider latency. MPC and the deterministic gate remain essential for safe, nonblocking control. The results support a simulation-based, two-time-scale Full agent; real-time use will require detector validation and evaluation under controller-model mismatch.

References

[1] H. Bagheri et al., "5G NR-V2X: Toward connected and cooperative autonomous driving," *IEEE Commun. Standards Mag.*, vol. 5, no. 1, pp. 48–54, Mar. 2021.

[2] J. Lee et al., "Overview of 5G-NR-V2X system and analysis methodology of communication performance," *J. IKEEE*, vol. 28, no. 3, pp. 310–320, 2024.

[3] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 2731–2735.

[4] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.

[5] X. Xu et al., "Enhanced velocity-adaptive scheme: Joint fair access and age of information optimization in vehicular networks," *IEEE Trans. Mobile Comput.*, vol. 25, no. 3, pp. 3488–3505, Mar. 2026.

[6] X. Xu et al., "Velocity-adaptive access scheme for semantic-aware vehicular networks: Joint fairness and AoI optimization," *IEEE Trans. Mobile Comput.*, vol. 25, no. 8, pp. 12205–12221, Aug. 2026.

[7] M. Ji et al., "Joint optimization of trajectory control, resource allocation, and task offloading for multi-UAV-assisted IoV," *IEEE Trans. Mobile Comput.*, vol. 25, no. 9, pp. 15342–15359, Sep. 2026.

[8] A. Baiocchi and I. Turcanu, "Age of information of one-hop broadcast communications in a CSMA network," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 294–298, Jan. 2021.

[9] A. Rolich, I. Turcanu, and A. Baiocchi, "AoI-aware and persistence-driven congestion control in 5G NR-V2X sidelink communications," in *Proc. MedComNet*, Jun. 2024, pp. 1–4.

[10] S. Cho, Y. Kim, and Y. Han, "Analysis of control overhead reduction by semi-persistent scheduling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 726–730, May 2020.

[11] M. Muhammad Saad, M. Ashar Tariq, Md. Mahmudul Islam, M. Toaha Raza Khan, J. Seo, and D. Kim, "Enhanced semi-persistent scheduling (e-SPS) for aperiodic traffic in NR-V2X," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, Feb. 2022, pp. 171–175.

[12] M. M. Saad, M. A. Tariq, J. Seo, M. Ajmal, and D. Kim, "Age-of-information aware intelligent MAC for congestion control in NR-V2X," in *Proc. 14th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2023, pp. 265–270.

[13] A. Rolich, I. Turcanu, A. Vinel, and A. Baiocchi, "Understanding the impact of persistence and propagation on the age of information of broadcast traffic in 5G NR-V2X sidelink communications," *Comput. Netw.*, vol. 248, Jun. 2024, Art. no. 110503.

[14] N. C. Luong et al., "Applications of deep reinforcement learning in communications and networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3133–3174, 4th Quart., 2019.

[15] M. Ji et al., "Graph neural networks and deep reinforcement learning-based resource allocation for V2X communications," *IEEE Internet Things J.*, vol. 12, no. 4, pp. 3613–3628, Feb. 2025.

[16] Z. Huan, J. Sun, Z. Chen, Z. Zhang, X. Sun, and Z. Xiao, "Resource allocation in V2X networks: A double deep Q-network approach with graph neural networks," *Comput., Mater. Continua*, vol. 84, no. 3, pp. 5427–5443, 2025.

[17] X. Chen et al., "Age of information aware radio resource management in vehicular networks: A proactive deep reinforcement learning perspective," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2268–2281, Apr. 2020.

[18] X. Gu, J. Peng, L. Cai, X. Zhang, and Z. Huang, "Markov analysis of C-V2X resource reservation for vehicle platooning," in *Proc. IEEE 95th Veh. Technol. Conf. (VTC-Spring)*, Jun. 2022, pp. 1–5, doi: 10.1109/VTC2022-Spring54318.2022.9860899.

Authorized licensed use limited to: Southeast University. Downloaded on October 30,2026 at 12:09:54 UTC from IEEE Xplore. Restrictions apply.