Dynamic Pricing and Incentivization in Multimodal Transportation Networks using Deep Reinforcement Learning

Khadidja Kadem*[1], Mostafa Ameli[1], Carlos Lima Azevedo[2], Mahdi Zargayouna[1], and Latifa Oukhellou[1]

[1]COSYS-GRETTIA, Gustave Eiffel University, France
[2]Department of Technology, Management and Economics, Technical University of Denmark, Denmark

Short summary

This paper proposes a multi-agent deep reinforcement learning (DRL) framework for coordinating dynamic pricing and incentivization strategies in multimodal transportation networks. Two competing DRL agents are deployed: a shared mobility service provider that maximizes revenue through adaptive fare adjustment, and a public authority that allocates spatio-temporal transit subsidies to improve equity, reduce emissions, and enhance system efficiency. Unlike prior work that treats pricing and incentivization separately, the proposed multi-agent DRL formulation captures their dynamic interdependencies within a multimodal network, enabling more balanced and system-level optimal outcomes. Experiments on the Sioux Falls benchmark network demonstrate that dynamic incentivization reduces commuters' costs by around 20%, cuts emissions by approximately 10%, and nearly doubles public transport profit during the morning peak, while the joint dynamic pricing-incentivization strategy successfully balances conflicting objectives between private and public actors. The proposed approach can support the analysis of sustainable and equitable multimodal mobility strategies.
Keywords: deep reinforcement learning, incentivization, intermodality, pricing, shared mobility, transport equity

1 Introduction

The rapid growth of urban populations is intensifying pressure on cities to develop more sustainable and space-efficient mobility systems. Private vehicles continue to dominate urban travel, contributing disproportionately to traffic congestion, air pollution, and greenhouse gas emissions. In this context, shared mobility services (SMSs), such as ridesharing and carpooling, have emerged as a promising complement to public transport (PT), offering higher vehicle occupancy, flexible first/last-mile connectivity, and reduced car dependency. Yet, their integration into existing multimodal networks brings complex trade-offs: SMSs can attract users away from PT, concentrate demand in high-density zones, and exacerbate spatial and social inequities if left unregulated (Oh et al., 2020). Understanding and managing these effects requires a modeling and optimization framework that captures the full complexity of multimodal systems and reconciles the conflicting objectives of the stakeholders.

A significant body of research has developed models to capture traveler behavior and system performance in multimodal networks. Static equilibrium-based approaches offer a tractable and unified formulation of mode choice, user heterogeneity, and intermodality (Kadem et al., 2025). However, their static nature is a fundamental limitation, as they yield snapshots of equilibrium states and cannot capture the time-varying dynamics of demand, congestion, and service availability that characterize real urban networks during peak periods. Dynamic models address this gap through agent-based (Oh et al., 2020) or macroscopic simulation frameworks (Lee et al., 2025). These tools offer greater realism, but often at the cost of complexity: many remain limited in their treatment of intermodality or require substantial modeling effort to adapt to new policy contexts. There is therefore a need for dynamic frameworks that are both comprehensive and tractable enough to support systematic policy evaluation.

Beyond traveler behavior, the strategic decisions of private SMSs operators constitute a critical driver of system dynamics. Ridesharing platforms continuously adjust their fares in response to

supply-demand imbalances to maximize revenue. A rich body of research has studied these dynamic pricing strategies using analytical models (Ma et al., 2020) and, more recently, reinforcement learning (RL) approaches (Ge et al., 2025). However, these studies are often developed in isolation from the broader multimodal system, neglecting the effect that pricing decisions have on modal choice, network congestion, and PT attractiveness. Moreover, equity considerations, such as how pricing strategies affect the costs of different groups or spatial areas, remain largely underexplored in this literature.

In this context, public authorities must rely on policy instruments to steer the system toward socially desirable outcomes. Incentivization, in the form of subsidies, has been proposed as one such instrument, with studies exploring its use to shift departure times (Li et al., 2018), promote vehicle sharing (Sun et al., 2022), and improve PT ridership (Gao & Li, 2024). Some works have begun to leverage RL for adaptive incentive design (Chen & Wang, 2025), recognizing that static subsidy schemes cannot respond to evolving demand and congestion. However, these approaches often treat incentivization independently of the concurrent strategic behavior of a profit-driven SMS provider, and the implications of such policies on equity remain insufficiently studied.

Addressing this gap is the central objective of this paper. We propose a dual-agent deep reinforcement learning framework in which a ridesharing provider and a public authority simultaneously learn spatio-temporal strategies within a shared multimodal simulation environment. The provider dynamically adjusts fares to maximize profit, while the authority deploys PT incentives to optimize system performance across three dimensions: efficiency (total generalized cost), environmental impact (CO emissions), and equity (distribution of travel cost benefits across user classes and geographic zones). By explicitly representing each stakeholder as a learning agent that responds to the other's decisions, our framework accounts for strategic interdependencies that prior work has overlooked, and enables a rigorous evaluation of the trade-offs between private profitability and public objectives. To the best of our knowledge, this is the first study to jointly optimize dynamic pricing and incentivization under a multi-agent DRL paradigm, while simultaneously targeting efficiency, emissions, and equity.

2 Methodology

The proposed framework models the multimodal transportation system as a multi-agent deep reinforcement learning problem in which two learning agents, a public authority and a ridesharing (RS) provider, interact within a shared simulation environment. As illustrated in Figure 1, both agents observe the current state of the multimodal system and select their respective actions, which directly affect the fares faced by travelers, thereby influencing travel demand, modal choices, and overall network conditions. The resulting system state is then fed back into the learning process of both agents.

In this study, we assume that both agents operate on the same temporal scale, which allows decisions to be made concurrently and ensures a consistent and tractable representation of their interaction. This multi-agent simulation-based framework therefore supports the evaluation of short-term policy interventions while accounting for the dynamic feedback that governs travel behavior and system performance.

The following sections describe each DRL agent's decision problem, before detailing the simulation environment in which they interact.

Public Authorities

The public authority aims to improve system-level performance by offering spatio-temporal fare reductions on PT. Its decision problem is defined as follows.

- **State:** At each timestep τ, the agent observes the level of PT demand departing from and arriving at each node of the network, the current time of day τ, and the remaining subsidy budget. This state representation gives the agent visibility over where and when PT demand is concentrated, enabling spatially targeted incentive allocation.
- **Action:** The action space consists of origin-destination fare reductions $\delta^{(i,j)}(\tau)$ applied to PT trips. To keep the problem tractable, the OD incentive is decomposed into an origin-based component $\delta^i_o(\tau)$ and a destination-based component $\delta^j_d(\tau)$), with the resulting OD value computed as their average:

$$\delta^{(i,j)}(\tau) = \frac{1}{2}(\delta^i_o(\tau) + \delta^j_d(\tau)) \qquad \delta^i_o(\tau), \delta^j_d(\tau) \in I \tag{1}$$

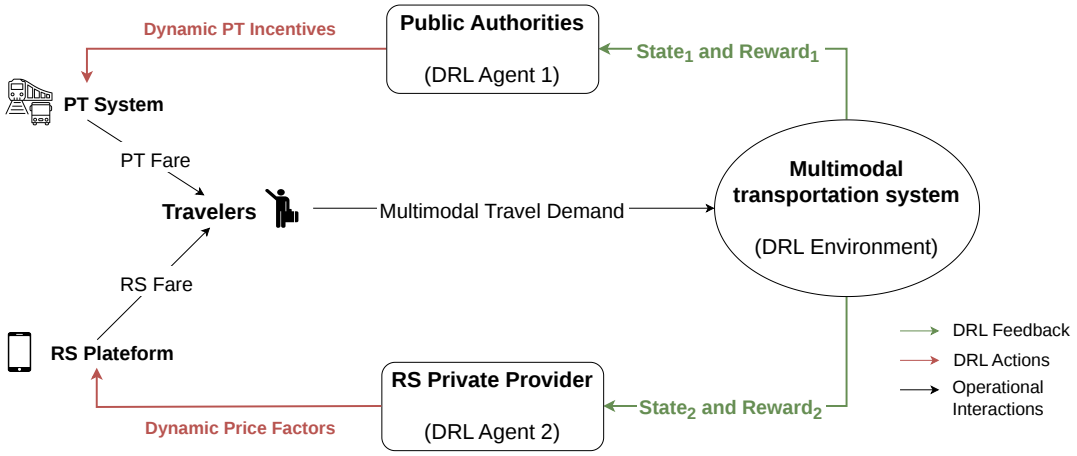


Figure 1: Dual-Agent DRL architecture for the joint dynamic pricing and incentivization strategy

where I is the set of predefined incentives. The actual fare paid by a PT passenger on OD pair (i, j) at time τ is then $F_{PT}^{(i,j)}(\tau) = b_{PT} - \delta^{(i,j)}(\tau)$, where b_{PT} is the base PT fare.

- **Reward:** The objective of authorities is to optimize overall system performance. For this, we define the system performance with three indicators:
  - **Total Generalized Cost (TGC)** as a metric of efficiency. It is measured by aggregating individual travel costs across all travelers in the system.
  - **System Emission Rate (SER)**, computed through a link-based CO emission model that relates emission rates non-linearly to link travel time and length.
  - **Gini index (G)** as a metric of inequity. It measures how evenly the travel cost benefits of the policy are distributed across user classes and geographic zones. A low value indicates a homogeneous distribution of benefits, while a high value signals a concentration of gains in particular groups or areas.

We adopt in this study a single-objective training strategy and learn three separate DRL policies for the public authority. Training separate agents per objective allows us to isolate the effect of each public goal on the system's behavior and on the resulting interaction with the ridesharing provider.

Ridesharing Provider

The ridesharing provider is modeled as a profit-maximizing private operator that dynamically adjusts fares in response to supply and demand across the network. Its decision problem is defined as follows.

- **State:** At each timestep, the agent observes the level of ridesharing demand departing from and arriving at each node, the number of available vehicles per node, and the time of day. This gives the agent the real-time demand-supply picture needed to identify zones where price adjustments can increase revenue or attract additional riders.
- **Action:** The action space consists of origin-destination price factors $\lambda^{(i,j)}(\tau)$ that scale the base ridesharing fare. Following the same dimensionality reduction as the authority agent, the OD factor is expressed as the average of an origin-based $\lambda_o^i(\tau) \in]0, 2]$ and a destination-based $\lambda_d^j(\tau) \in]0, 2]$ multipliers:

$$\lambda^{(i,j)}(\tau) = \frac{1}{2}(\lambda_o^i(\tau) + \lambda_d^j(\tau)) \tag{2}$$

The fare paid by a ridesharing passenger using path p on OD pair (i, j) is then $F_{p,RS}^{(i,j)}(\tau) = \lambda^{(i,j)}(\tau) \times b_{RS} \times L_p$ where b_{RS} is the base fare per unit distance and L_p is the path length.

- **Reward:** The agent's reward is its instantaneous profit, defined as fare revenues from served ridesharing trips minus vehicle operating costs. This reward directly encourages the agent to balance fare levels against demand: raising prices increases per-trip revenue but risks losing riders to PT or private vehicles, while lowering prices attracts volume at reduced margins.

Multimodal Simulation Environment

The simulation environment constitutes the shared world in which both agents act. It is composed of three coupled modules operating within a rolling horizon scheme. The planning horizon is discretized into time intervals of equal length ΔT. At each timestep τ, the system evolves through the sequence of operations illustrated in Figure 2 and described below, after which the agents receive their rewards and the horizon advances.

- The **traffic assignment** allocates travel demand by solving a multi-class user equilibrium problem. Travelers are cost-minimizing agents choosing their routes and modes simultaneously. Available modes are: private vehicle, bus, metro/railway, carpooling (driver and passenger), ridesharing, walking, and biking, as well as intermodal combinations. Their generalized cost integrates trip fare, inconvenience factor, as well as travel, waiting, and service time, weighted by a class-specific value of time reflecting income-based heterogeneity. Travelers' choices continuously respond to both the provider's pricing and the authority's incentives, since fares are updated at each timestep.
- The **SMSs matching** translates the SMSs demand flows produced by the assignment into vehicle schedules. For each available vehicle, it finds the optimal sequence of pickup and drop-off stops satisfying passengers' time windows and vehicle capacity, through a two-phase procedure combining a greedy passenger-vehicle assignment with a parallelized branch-and-bound scheduler. Unmatched requests are returned to the assignment module and redirected to alternative modes.
- The **traffic simulator** executes all vehicle movements, namely matched SMSs trips, PT services, private vehicles, and soft modes, using a trip-based macroscopic fundamental diagram (MFD) model. PT is modeled at an aggregate frequency-based level. Metro services operate on a dedicated layer at constant speed, while buses share the road network and are subject to MFD-based congestion. The simulator outputs updated average speeds and waiting times, which feed back into the assignment module.

Convergence is achieved when speeds and waiting times vary by less than a defined threshold between consecutive iterations, at which point the DRL agents receive their rewards and the horizon advances to $\tau + \Delta T$.

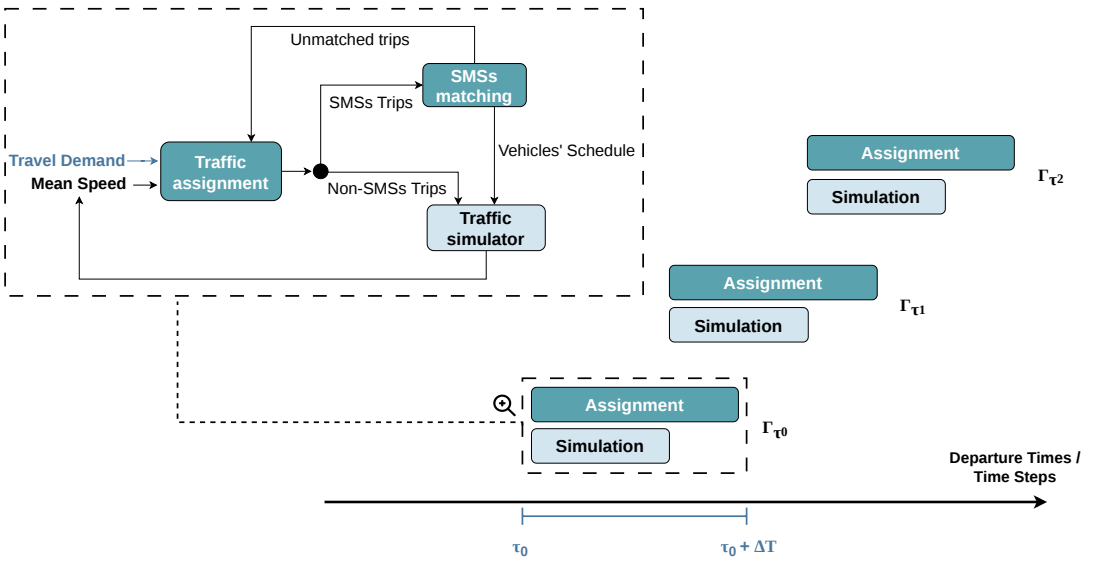


Figure 2: Rolling horizon approach for multimodal systems

A critical property of this approach is that both DRL agents act concurrently and fully share the

simulation environment. Thus, each agent's reward depends not only on its own actions but also on the strategy adopted by the other. Neither agent can therefore optimize its decisions in isolation.

3 Results and discussion

Experiments are conducted on the Sioux Falls benchmark network, illustrated in Figure 3. Three user classes (low, medium, high VoT) are spatially distributed based on socioeconomic data. Departure times are sampled stochastically to reflect a realistic morning peak profile.

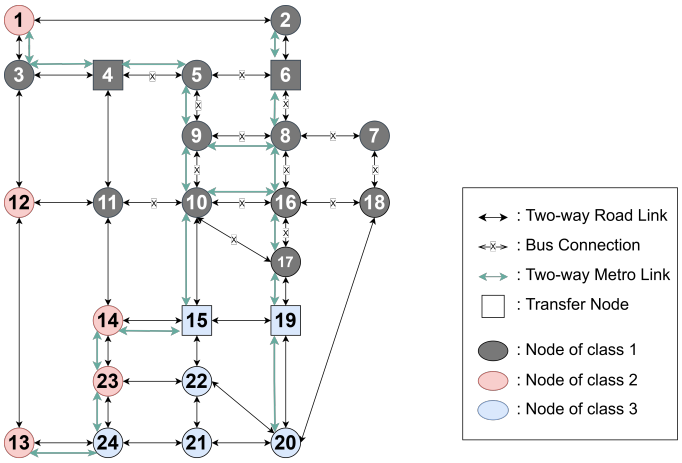
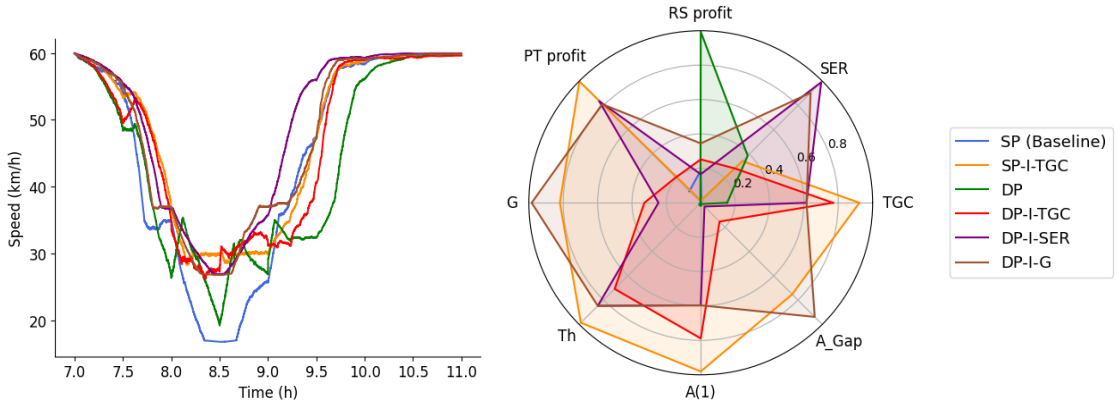


Figure 3: Sioux Falls network with PT infrastructure and heterogeneous user classes

We evaluate a set of scenarios designed to assess how the two RL agents interact and what system-level trade-offs emerge from their joint operation. The scenarios are defined as follows:

- **SP:** static ridesharing pricing with no PT incentives. This represents the reference scenario against which all others are compared.
- **SP-I-TGC:** static ridesharing pricing with RL-based PT incentivization targeting TGC. This isolates the effect of adaptive public policy in the absence of a strategic provider.
- **DP:** dynamic RL-based pricing with no PT incentives. This isolates the effect of a profit-maximizing provider operating without regulation.
- **DP-I-TGC:** joint dynamic pricing and PT incentivization for TGC.
- **DP-I-SER:** joint dynamic pricing and PT incentivization for SER.
- **DP-I-G:** joint dynamic pricing and PT incentivization for G.

Figure 4a presents the evolution of average speed across scenarios. Figure 4b summarizes performance across six dimensions: TGC, SER, ridesharing provider profit, PT profit, and equity assessed through the Gini index and other well-established inequality metrics, namely the Theil coefficient (Th), Atkinson index (A(1)), and the accessibility gap between income groups (A_Gap). Each metric is normalized so that higher values on the chart correspond to better performance.
The DP scenario confirms that a profit-maximizing provider can generate efficiency and emission gains relative to the baseline, primarily by increasing vehicle sharing and redistributing demand across the network. However, these gains come at a high cost to equity and PT attractiveness. Moreover, the speed profile reveals noticeable oscillations throughout the peak period, a direct consequence of the provider's reactive pricing behavior.
The introduction of PT incentivization alongside dynamic pricing consistently stabilizes these oscillations and shifts the system toward more socially desirable outcomes, though the nature of the improvement depends on the authority's objective. The equity-oriented configuration (DP-I-G) produces the most balanced outcome across all evaluated dimensions, by explicitly minimizing disparities in the distribution of travel cost benefits. Thus, PT profit increases by 74% and equity improves compared to the baseline (SP), while ridesharing still captures a significant share of demand and maintains profitability.
These results demonstrate that the proposed framework captures foresighted dynamic stakeholder interactions and enables meaningful reconciliation of private and public objectives.



(a) (b)

Figure 4: (a): Temporal evolution of average network speed. (b): Comparative analysis of the performance metrics

4 Conclusion

This paper proposes a multi-agent deep reinforcement learning framework for jointly optimizing dynamic pricing and incentivization in multimodal transportation networks. By modeling a profit-driven ridesharing provider and a system-oriented public authority as interacting learning agents within a macroscopic simulation, the framework captures key strategic foresighted interdependencies often neglected in prior work.

Experiments on the Sioux Falls network highlight three main findings. First, dynamic pricing alone increases provider revenue, slightly reduces congestion and emissions, but worsens equity and reduces PT attractiveness. Second, combining pricing with incentivization mitigates these adverse effects. Third, an equity-oriented incentivization strategy yields the most balanced performance across all dimensions. These findings suggest that equity-oriented incentivization has the potential to support sustainable system optimization, although outcomes remain dependent on network characteristics and demand patterns.

Several directions remain open for future research. Scaling the framework to larger and more realistic urban networks will require advances in computational efficiency. Moreover, incorporating vehicle repositioning and non-compliant driver behavior would add important operational realism. Finally, extending the authority agent toward multi-objective learning, allowing it to self-balance efficiency, emissions, and equity, would bring the framework closer to deployment as a practical decision-support tool for sustainable urban mobility planning.

References

Chen, Y., & Wang, S. (2025, August). Optimization of dynamic incentive strategies for public transportation based on reinforcement learning and network synergy effect. *Scientific Reports*, *15*(1), 30695. Retrieved 2025-12-05, from `https://www.nature.com/articles/s41598-025-16632-y` (Publisher: Nature Publishing Group) doi: 10.1038/s41598-025-16632-y

Gao, J., & Li, S. (2024, May). Regulating for-hire autonomous vehicles for an equitable multimodal transportation network. *Transportation Research Part B: Methodological*, *183*, 102925. Retrieved 2024-10-15, from `https://www.sciencedirect.com/science/article/pii/S0191261524000493` doi: 10.1016/j.trb.2024.102925

Ge, S., Zhou, X., & Qiu, T. (2025, May). R2Pricing: A MARL-Based Pricing Strategy to Maximize Revenue in MoD Systems With Ridesharing and Repositioning. *IEEE Transactions on Mobile Computing*, *24*(5), 3552–3566. Retrieved 2025-04-23, from `https://ieeexplore.ieee.org/document/10787072/` doi: 10.1109/TMC.2024.3514124

Kadem, K., Ameli, M., Zargayouna, M., & Oukhellou, L. (2025, August). *Multimodal urban transportation network equilibrium including intermodality and shared mobility services.* arXiv.

Retrieved 2025-12-19, from http://arxiv.org/abs/2402.00735 (arXiv:2402.00735 [math]) doi: 10.48550/arXiv.2402.00735

Lee, E., Slama, R., & Leclercq, L. (2025, October). Assessing how ride-hailing rebalancing strategies improve the resilience of multi-modal transportation systems. *Transportation Research Part C: Emerging Technologies*, *179*, 105300. Retrieved 2025-09-23, from https://www.sciencedirect.com/science/article/pii/S0968090X25003043 doi: 10.1016/j.trc.2025.105300

Li, C., Ma, J., Luan, T. H., Zhou, X., & Xiong, L. (2018, October). An incentive-based optimizing strategy of service frequency for an urban rail transit system. *Transportation Research Part E: Logistics and Transportation Review*, *118*, 106–122. Retrieved 2025-07-18, from https://linkinghub.elsevier.com/retrieve/pii/S1366554517307433 (Publisher: Elsevier BV) doi: 10.1016/j.tre.2018.07.005

Ma, J., Xu, M., Meng, Q., & Cheng, L. (2020, April). Ridesharing user equilibrium problem under OD-based surge pricing strategy. *Transportation Research Part B: Methodological*, *134*, 1–24. Retrieved 2024-03-14, from https://linkinghub.elsevier.com/retrieve/pii/S0191261519303832 doi: 10.1016/j.trb.2020.02.001

Oh, S., Seshadri, R., Azevedo, C. L., Kumar, N., Basak, K., & Ben-Akiva, M. (2020, August). Assessing the impacts of automated mobility-on-demand through agent-based simulation: A study of Singapore. *Transportation Research Part A: Policy and Practice*, *138*, 367–388. Retrieved 2024-05-13, from https://linkinghub.elsevier.com/retrieve/pii/S0965856420306133 doi: 10.1016/j.tra.2020.06.004

Sun, Y., Chen, S., & Guo, Q. (2022, July). Evaluating the Environmental Benefits of Personalized Travel Incentives in Dynamic Carpooling. *KSCE Journal of Civil Engineering*, *26*(7), 3082–3093. Retrieved 2025-01-16, from https://linkinghub.elsevier.com/retrieve/pii/S1226798824015228 doi: 10.1007/s12205-022-1568-1