Joint Optimization of Pricing and Service Type in Ride-Hailing using Multi-agent Reinforcement Learning

Tao Peng*[1], Jie Gao[1], and Oded Cats[1]

[1]Department of Transport & Planning, Delft University of Technology, the Netherlands

Short summary

Ride-hailing systems often offer both private hailing and pooling ride services, and platform performance depends on both how drivers are allocated across these services and how each service is priced. However, most existing studies treat these decisions separately or focus on a single service type, either private hailing or pooling. This paper proposes a multi-agent reinforcement learning framework to jointly optimize pricing and service type in ride-hailing systems under endogenous rider choice. Specifically, we formulate the problem as a multi-agent Markov Decision Process and train a pricing agent together with a service type agent using Proximal Policy Optimization in a realistic simulation environment. To ensure scalability across different fleet sizes, we employ a parameter-sharing architecture among service agents. Experiments on a realistic Delft road network demonstrate that the proposed framework learns stable and effective policies, consistently outperforming several rule-based baselines. The results highlight the substantial benefits of jointly optimizing pricing and service type in ride-hailing systems.
Keywords: Multi-agent reinforcement learning, Joint Optimization, Ride-hailing, Shared mobility

1 Introduction

Ride-hailing systems often offer both private hailing and pooling services, and the way these services are designed has an important influence on system performance. A suitable combination of private hailing and pooling services can improve vehicle utilization, increase served demand, and enhance platform profit by better matching limited vehicle capacity to heterogeneous travel demand. However, riders do not choose between private hailing and pooling based on service type alone. Their willingness to accept an offered service depends on its fare, waiting time, and detour(Alonso-González et al., 2021; Kang et al., 2021; Liu et al., 2019; Li, 2023). These factors determine whether riders choose the private hailing service, the pooling service, or opt out of the system. Since the platform directly controls the service type and price, these two decisions must be designed jointly to improve overall system performance, including platform profit, average waiting time, and detours.

A growing body of work has examined this design problem from a static perspective. For example, Jacob & Roet-Green (2021) study pricing menus for private hailing and pooling services under stochastic demand and heterogeneous passenger valuations, while Bahrami et al. (2022) analyze whether platforms should offer private hailing-only, pooling-only, or mixed services. Other studies incorporate both pricing and allocation decisions under specific assumptions on demand and user heterogeneity (Wu et al., 2025; Zhong et al., 2023; K. Zhang & Nie, 2021). Although these works provide important insights into the interaction between pricing and service type, they are developed under static or equilibrium settings. In practice, however, ride-hailing systems evolve continuously over time, and current decisions affect future rider acceptance, fleet availability, and matching opportunities. This makes the problem inherently sequential, and dynamic service type and pricing decisions become critical for long-term performance.

Reinforcement learning(RL) provides a natural framework for this setting. Recent RL-based studies have applied RL to a variety of operational decisions in ride-hailing systems. In particular, several works focus on dynamic pricing (Chen et al., 2021; Shi et al., 2022; Lei & Ukkusuri, 2023), while others jointly optimize pricing with dispatching, repositioning, or scheduling (Huang et al., 2022; Z. Zhang et al., 2024; Ge et al., 2024). Although these approaches better capture interactions across multiple decision components, several limitations remain. First, service type is not modeled as a dynamic decision variable, which limits the platform's ability to reallocate fleet capacity between private hailing and pooling services in response to changing conditions. Second, rider behavior is modeled in a simplified manner (e.g., via reduced-form acceptance functions), limiting the ability to capture heterogeneous sensitivities to key service attributes. Third, jointly optimizing multiple decision dimensions can lead to substantial scalability challenges, especially when service type decisions must be made at the driver level for large fleets.

To address these limitations, we propose a multi-agent reinforcement learning framework for the joint dynamic optimization of service type and pricing in ride-hailing systems. The framework coordinates a service type agent and a pricing agent to jointly determine service allocation and pricing over time. To explicitly model rider choice, we incorporate a utility-based choice model in which riders respond to fare, waiting time, and detour. To scale service type decisions to varying fleet sizes, we adopt a parameter-sharing architecture and train the framework using Proximal Policy Optimization in a realistic simulation environment. By jointly optimizing service type and pricing under this setting, the proposed approach improves rider acceptance, increases served demand, and enhances platform profit through a more efficient use of vehicle capacity.
The main contributions of this work are as follows, we:

- formulate the joint optimization of pricing and service type as a Markov Decision Process (MDP), where the key decisions include setting prices for each service type and dynamically assigning vehicles to different service types.
- design a deep multi-agent reinforcement learning (MARL) framework based on the proposed MDP formulation. To ensure scalability, we adopt a parameter-sharing mechanism to mitigate the explosion of policy parameters and action space.
- develop an efficient and realistic simulator that models the operation of ride-hailing and ride-pooling services simultaneously. Built on a real-world urban road network, the simulator includes key functionalities such as request generation and spatial matching. Moreover, it incorporates a realistic rider choice model that captures sensitivity to differentiated service attributes, enabling a more reliable evaluation of decision policies.
- evaluate the proposed approach through extensive simulation experiments on a real-world urban network, comparing it against seven rule-based baselines. The results demonstrate consistent improvements in platform revenue, number of served requests, and fleet utilization compared to conventional strategies.

2 A MULTI-AGENT REINFORCEMENT LEARNING FRAMEWORK FOR JOINT SERVICE TYPE AND PRICING OPTIMIZATION

Problem Description

We consider a ride-hailing system in which a platform matches rider requests with available drivers over time. The platform provides two types of services: private hailing service, in which a driver serves an individual rider directly, and pooling service, in which a driver may serve multiple riders with compatible itineraries. At each time step, newly arrived riders submit requests characterized by their origins and destinations. The platform determines how private hailing and pooling services are offered, along with their corresponding prices. Given these alternatives, riders choose among private hailing, pooling, and the opt-out option based on a utility-based model that depends on service attributes such as price, expected waiting time, and detour. Based on realized rider choices, the system performs service-specific matching between riders and drivers. These matching outcomes determine immediate system performance, including served demand and platform revenue, as well as influence future service opportunities. The platform therefore aims to dynamically determine service type and pricing decisions to optimize long-term system performance.

Multi-agent Markov Decision Process Modeling

We formulate the problem as a multi-agent Markov Decision Process (MAMDP), in which one pricing agent determines the prices of private hailing and pooling services, while a set of homogeneous service agents determines the service type assignment of available drivers. Formally, the MAMDP is denoted by the tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \gamma \rangle$. In this formulation, $\mathcal{S}$ denotes the state space, $\mathcal{A}$ is the joint action of price action $\mathcal{A}^p$ and service action $\mathcal{A}^s$, $\mathcal{T}$ is the transition probability function $\mathcal{T}(s_{t+1}|s_t, a_t)$, $\mathcal{R}$ is the shared reward function, and $\gamma \in [0, 1)$ is a discount factor. At each time step $t \in \{0, 1, \cdots, T\}$, the pricing agent and service agent observe the current state $s_t \in \mathcal{S}$, choose price action $a_t^p \in \mathcal{A}^p$ and service action $a_t^s \in \mathcal{A}^s$, and receive an immediate reward $\mathcal{R}(s_t, a_t, s_{t+1})$ after executing the actions and transitioning to the next state s_{t+1}. The optimization objective is to learn a joint policy π that maximizes the expected cumulative discounted reward over the time horizon:

$$\mathcal{J}(\pi) = \mathbb{E}_\pi \left[\sum_{t=0}^{\infty} \gamma \mathcal{R}(s_t, a_t, s_{t+1}) \right]$$

State: The system state at time t is defined as $s_t = (\mathbf{D}_t, \mathbf{R}_t, \mathbf{P}_t, h_t, \lambda_t)$, which captures both supply and demand information. $\mathbf{D}_t = \{d_i^t\}$ denotes the set of driver states, where each driver i is characterized by current location

$\mathbf{x}_i^t$, service mode $m_i^t \in \{\text{hailing}, \text{pooling}\}$, and current service type $\mathbf{c}_i^t$. For brevity, here we use "hailing" to denote private hailing in the mathematical notation, and the same abbreviations are used throughout the rest of the mathematical notations. $\mathbf{R}_t = \{u_j^t\}$ represents the set of rider states, where each rider j is described by origin $\mathbf{o}_j$, destination $\mathbf{d}_j$, service status τ_j^t, and value of time β_j. $\mathbf{P}_t = (p_t^{\text{hailing}}, p_t^{\text{pooling}})$ denotes the current unit prices for each service type. The temporal state $h_t \in [0, 24)$ encodes the time of day, and λ_t represents an aggregate demand intensity indicator, reflecting the overall level of rider demand at time t.

Action: The action of the pricing agent is represented by a vector of two continuous variables, $a_t^{\text{p}} = (p_t^{\text{hailing}}, p_t^{\text{pooling}}) \in \mathcal{A}^p \subseteq \mathbb{R}^2$, where p_t^{hailing} and p_t^{pooling} denote the unit prices for private hailing and pooling services, respectively. The action of the service agent is defined as a vector of binary variables, $a_t^{\text{s}} = \{s_t^i\}_{i=1}^{N_d} \in \mathcal{A}^s = \{0, 1\}^{N_d}$, where N_d is the number of drivers, and $s_t^i \in \{0, 1\}$ indicates whether driver i is assigned to the private hailing or pooling service at time step t.

Reward: The agents share a global reward signal designed to optimize platform efficiency while maintaining service quality. We define the reward as

$$r_t = \underbrace{\sum_{r \in \mathcal{R}_t} F_{r,t} + \omega_1 N_t^{\text{serve}} - \omega_2 N_t^{\text{idle}}}_{\text{primary reward}} - \underbrace{\omega_3 \frac{N_t^{\text{opt-out}}}{N_t}}_{\text{opt-out penalty}},$$

where $F_{r,t}$ is the trip fare, N_t^{serve} is the number of served riders, and N_t^{idle} is the number of idle drivers. While the primary reward captures revenue and service throughput, it does not explicitly account for unmet demand. To address this, we include a penalty based on the ratio of riders who choose the opt-out option such as travelling by buses or trains. Let $N_t^{\text{opt-out}}$ be the number of riders selecting the opt-out option, and N_t is the total number of riders. The coefficients ω_1, ω_2, and ω_3 balance the trade-offs between revenue, service throughput, and driver utilization.

Learning algorithm

We solve the proposed MAMDP using an actor-critic framework trained with Proximal Policy Optimization (PPO) (Schulman et al., 2017). The actor networks parameterize the policies (π) for sampling actions, while the critic networks estimate the state-value functions (V) used for advantage estimation. The workflow is shown in Figure 1. During training, the simulator generates trajectories by executing the joint action $a_t = (a_t^{\text{s}}, a_t^{\text{p}})$ sampled from the two policies π_s and π_p. For each transition $(s_t, a_t, r_t, s_{t+1}, \log \pi)$, where $\log \pi$ is the log-probability of action a_t under policy π at state s_t. The use of log-probabilities improves numerical stability by preventing underflow caused by extremely small probability values. The reward is stored in the experience buffer of both agents. Generalized Advantage Estimation (GAE) is used to compute advantages from the collected trajectories. Each agent then performs multiple epochs of PPO updates using mini batches sampled from the roll-out buffer.

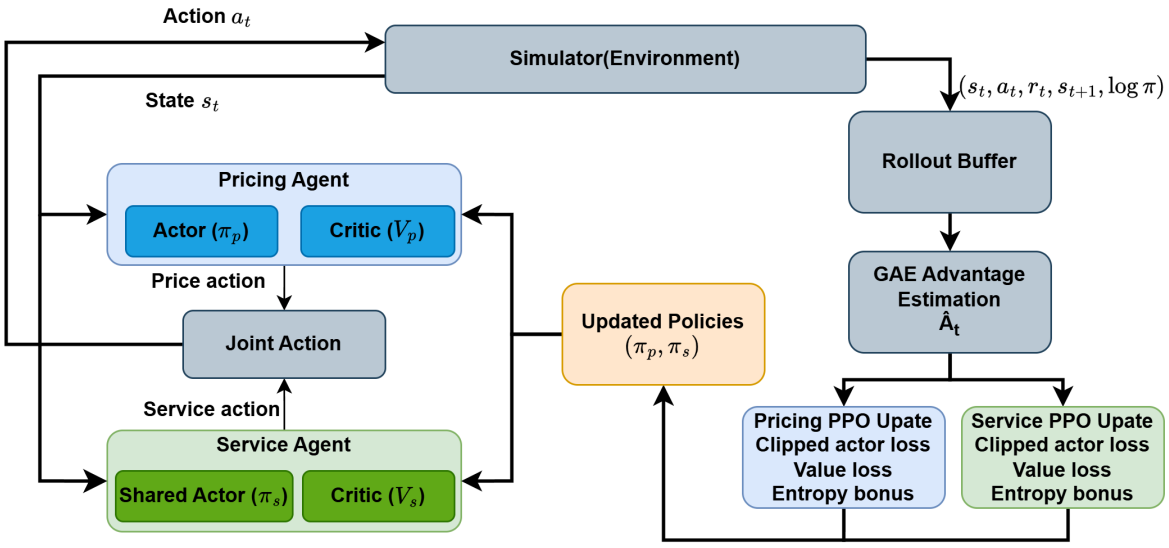


Figure 1: Algorithm structure of PPO

Although PPO provides the underlying policy optimization mechanism, scalability on the service side requires

additional architectural design. The service side decision is made at the driver level, which creates a large and variable action space when the fleet size changes. Learning an independent policy for each driver would therefore scale poorly and would not generalize across fleets of different size. To address this issue, we adopt a parameter-sharing (Hu et al., 1998) strategy in which a single policy network is shared across all drivers.
Let $\mathcal{N} = \{1, \ldots, N\}$ denote the set of drivers in the fleet. At each time step t, the system state is defined by a combination of global and local information. The global features vector $g_t \in \mathbb{R}^{d_g}$ represents system-wide information(e.g., time, demand distribution, and current price) shared by all drivers. The local features vector $l_{i,t} \in \mathbb{R}^{d_t}$ represents the specific state of driver i(e.g., location, service mode, load, and local surrounding rider density). The input representation for driver i, denoted as $x_{i,t}$ is the concatenation of these features:

$$x_{i,t} = [g_t, l_{i,t}] \in \mathbb{R}^{d_g + d_t}$$

We employ a shared policy network π parameterized by θ. This policy maps the input features to a Bernoulli probability $p_{i,t}$, representing the likelihood of vehicle i to take a service change action:

$$p_{i,t} = \pi(x_{i,t}; \theta)$$

where $p_{i,t} \in [0, 1]$. By sharing the parameters θ across all $i \in \mathcal{N}$, the model learns a universal decision rule that generalizes across the fleet regardless of its size.
Through this architecture and training protocol, the framework can jointly learn pricing and service type assignment policies that adapt to dynamic demand patterns while maintaining scalability for large driver fleets.

Simulator

To evaluate the proposed approach, we develop a simulator that captures the dynamics of private hailing and pooling services. Rider demand is generated at each time step according to a time-varying stochastic process with peak, standard, and off-peak periods.
To model rider behavior, we adopt a discrete choice framework based on price and travel time, including an opt-out option. For each rider r, the utility of service $k \in \{\text{hailing}, \text{pooling}, \text{opt-out}\}$ is defined as

$$U_{r,t}^k = V_r - \mathbb{E}[F_{r,t}^k] - \beta_r \mathbb{E}[T_{r,t}^k],$$

where, V_r denotes the intrinsic value of completing the trip , $\mathbb{E}[F_{r,t}^k]$ is the expected fare and $\mathbb{E}[T_{r,t}^k]$ is the expected total travel time (including both waiting time and in-vehicle time, with potential detours for pooling), converted into monetary disutility through the rider-specific value-of-time parameter β_r.
The expected waiting time is estimated based on local vehicle density, and pooling is assumed to incur 1.5 times longer waiting time due to additional pickup detours. Riders choose among alternatives according to a multinomial logit model:

$$P(c_r = m) = \frac{\exp(U_r^m)}{\sum_{k \in M} \exp(U_r^k)}, \quad m \in M = \{\text{hailing}, \text{pooling}, \text{opt-out}\}.$$

On the supply side, we consider a fixed fleet of drivers. Private hailing requests are matched to the nearest available driver, while pooling requests must satisfy vehicle capacity and route constraints. Vehicles travel along shortest paths on the road network. At each time step, the platform determines the price for each service type and and assigns a service type to each driver, after which rider choices, matching, and trip execution are realized, and the system transitions to the next state.

3 Model demonstration

To evaluate the proposed framework, we conduct experiments using a Delft (the Netherlands) city map extracted from OpenStreetMap. As shown in Figure 2, the road network covers the central area of Delft, containing 2424 nodes and 5707 edges, with a total length of approximately 400 km. Riders are generated at different nodes across the city, while drivers move along the road network to pick up and drop off passengers. Vehicle speed is set to 30 km/h, which reflects a reasonable average speed for urban traffic conditions.

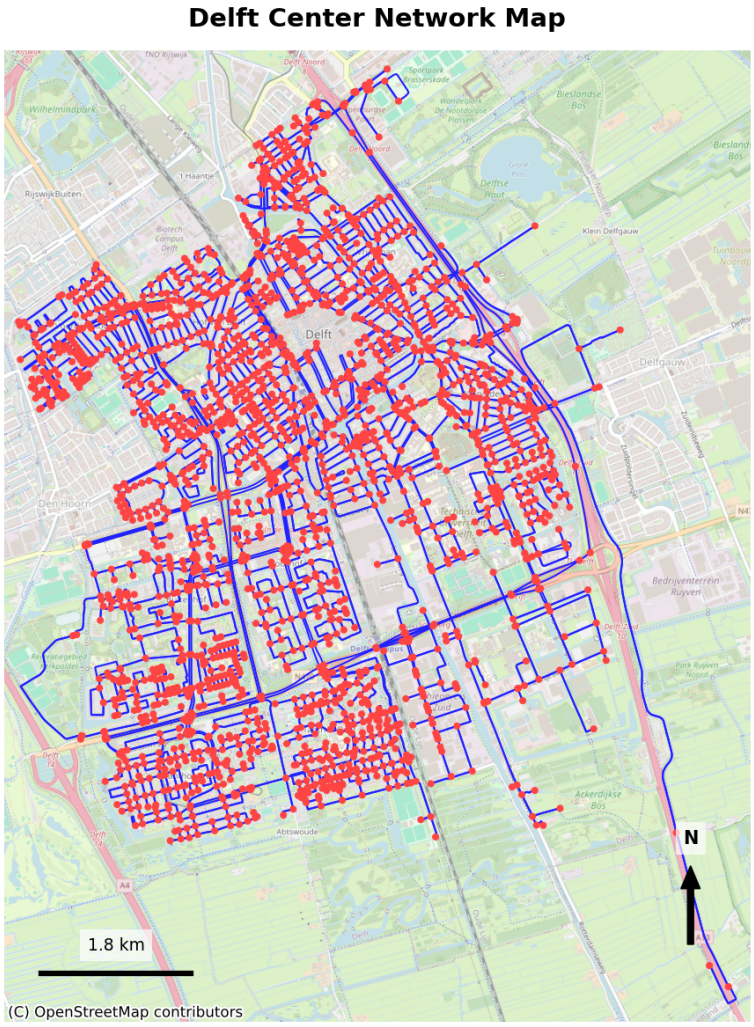
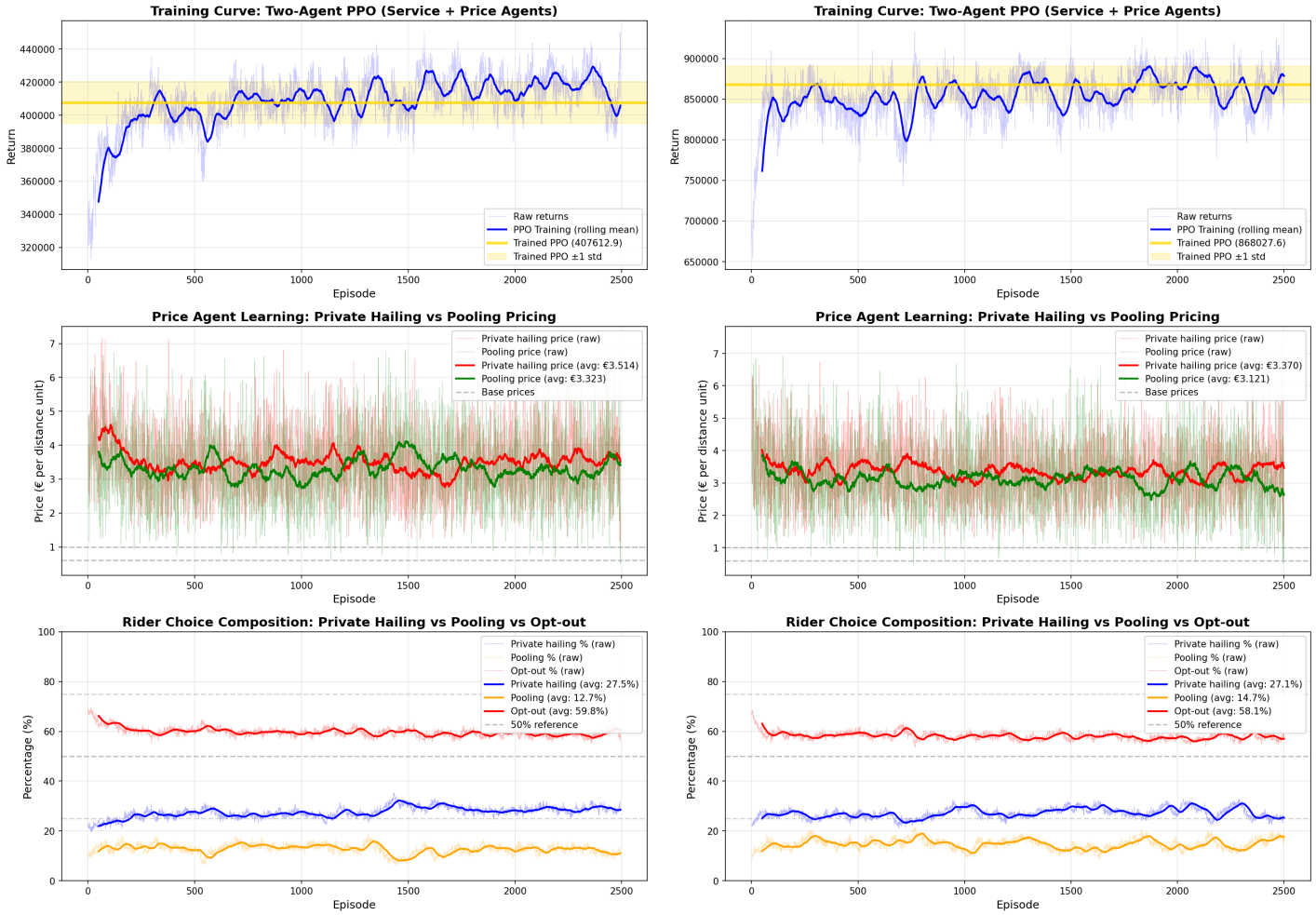


Figure 2: Delft network map, containing 2424 nodes and 5707 edges.

For parameter settings such as riders' value of time, public transit base fare, and public transit fare per kilometer, we follow the configuration used in K. Zhang & Nie (2021). The speed of public transit is assumed to be 1.5 times slower than that of private vehicles. In addition, transit headway times are assumed to vary according to peak and off-peak demand patterns. During peak hours, public transit services operate more frequently, resulting in shorter waiting times, whereas waiting times increase during off-peak periods. Through this design, we model public transit as a strong competitor to ride-hailing services, reflecting realistic urban mobility conditions.

Training curve

In the experimental setup, we consider 100 drivers and up to 150 riders. At the start of each episode, drivers are randomly distributed across the city with randomly assigned service modes. At each time step, new riders arrive: 35 during peak hours, 20 during off-peak hours, and 28 during standard periods. These settings create varying demand conditions while maintaining sufficient system pressure to evaluate adaptive pricing and service assignment. The simulation horizon is 500 time steps.
We train the framework for 2500 episodes and report the evolution of system return, pricing, and rider choice in Figure 3a. The system return increases steadily from around 250 000 to over 400 000 within 300 episodes and stabilizes near this level, peaking at approximately 450 000, indicating a substantial performance improvement. The learned prices initially decrease, with private hailing priced slightly higher than pooling. Over time, the two converge to similar levels, with the average private hailing price remaining slightly higher than pooling (3.528 € vs. 3.336 €) in the final 100 episodes. However, the pooling price occasionally exceeds the Private hailing price. This occurs because the learned policy samples private hailing and pooling prices independently, without imposing a constraint that pooling must be cheaper than private hailing. Stochastic policy noise and ongoing exploration therefore lead to episodes in which the realized pooling price exceeds the private hailing price, even though the average private hailing price remains higher at the end of training.
Rider choice dynamics show that the opt-out option remains dominant but decreases from about 75% to 60%.



(a) Training results with 100 drivers and 150 riders. (b) Training results with 200 drivers and 300 riders.

Figure 3: For each image, the top panel shows system return during training, including the raw return (light blue), rolling mean (dark blue), and the average return over the last 100 episodes (yellow). The middle panel presents the evolution of service prices, including private hailing (light red for raw and red for rolling mean) and pooling (light blue for raw and blue for rolling mean). The bottom panel shows rider choice proportions, including private hailing (yellow), pooling (blue), and opt-out option (red), where lighter colors indicate raw values and darker colors indicate rolling averages.

Meanwhile, the share of private hailing increases from less than 20% to around 30%, while pooling remains stable at approximately 15%, suggesting improved service attractiveness.
We further evaluate a setting with 200 drivers and up to 300 riders (Figure 3b). In this case, convergence is faster (around 150 episodes), and the system achieves approximately twice the return. Both service prices are lower than 100-driver setting, and the rate of opt out choice decreases slightly, likely due to increased driver availability. Overall, these results show that the proposed framework learns effectively and converges to a stable policy. The learned policy improves system return and increases the share of riders choosing platform services, demonstrating the effectiveness of jointly optimizing pricing and service type.

Baseline comparisons

In this section, we compare the proposed approach with several rule-based strategies to evaluate the effectiveness of the joint learning framework. We consider a set of representative service type and pricing strategies. The service type strategies are described as follows:

- **Private hailing only**: all vehicles are assigned to private hailing service.
- **Private hailing dominates**: 70% of vehicles are assigned to private hailing service.
- **Pooling only**: all vehicles are assigned to pooling service.

- **Pooling dominates**: 70% of vehicles are assigned to pooling service.
- **Equal split**: 50% of vehicles are assigned to private hailing service and 50% to pooling service.
- **PPO-service**: the trained service agent from the proposed framework.

For pricing, the strategies are described as follows:

- **Supply-demand pricing**: where base prices (1.0 for private hailing and 0.6 for pooling) are adjusted by a factor of 1.5 under excess demand and 0.8 otherwise.
- **Time-of-day pricing**: where prices are scaled by 1.5 during peak hours and 0.8 during off-peak periods.
- **PPO-price**: the trained pricing agent from the proposed framework.

By combining these service assignment and pricing strategies, we construct seven baseline settings for comparison. We evaluate these baselines against the proposed framework, and record the mean return and total number of served riders. The results are reported in Table 1. When fixing the pricing strategy to PPO-price, performance varies significantly across service strategies. The *Private hailing dominated* strategy achieves the best performance (4424.49 return and 888.4 served riders), followed by *Private hailing only*. In contrast, *Pooling only* performs worst (821.16 return) and serves substantially fewer riders, indicating that excessive pooling supply can harm system efficiency when not aligned with demand.
When fixing the service strategy to PPO-service, rule-based pricing strategies lead to poor returns despite higher service levels. Both *Supply-demand* and *Time-of-day* pricing yield returns below 900, while serving over 1160 riders. This suggests that increased usage alone does not translate into higher revenue without effective pricing. The proposed framework, which jointly learns service assignment and pricing (*PPO-service and PPO-price*), achieves the best overall performance, with a return of 4960.75 and 1131 served riders. Compared to the best rule-based combination (*Private hailing dominated and PPO-price*), it improves return by approximately 12%. Notably, with the same PPO-service policy, switching from rule-based pricing (supply-demand or time-of-day) to the PPO-price policy increases the cumulative return by 5–8 times, at the cost of serving 3.4% fewer riders. These results highlight the importance of jointly optimizing service configuration and pricing to better coordinate supply and demand.

Table 1: Performance comparison of service and pricing strategies

Service strategy	**Pricing strategy**	**Return**	**Served riders**
Private hailing only	PPO-price	4320.70 ± 262.53	885.2
Private hailing dominated	PPO-price	4424.49 ± 379.46	888.4
Pooling only	PPO-price	821.16 ± 292.82	309.7
Pooling dominated	PPO-price	3400.35 ± 330.83	718.8
Balanced	PPO-price	3689.88 ± 429.31	763.9
PPO-service	Supply-demand	897.43 ± 175.02	1169.3
PPO-service	Time-of-day	685.35 ± 99.26	1170.2
PPO-service	**PPO-price**	4960.75 ± 265.55	1131

4 Conclusions

This study proposed a multi-agent reinforcement learning framework for the joint dynamic optimization of service type and pricing in ride-hailing systems with private hailing and pooling services. To address the scalability challenge arising from service type decisions over large fleets, the framework incorporates a parameter-sharing architecture for service agents. Experiments on a realistic simulator based on the Delft road network showed that the proposed method consistently outperformed rule-based baselines in served demand, vehicle utilization, and platform revenue. These results highlight the value of jointly optimizing service type and pricing in ride-hailing systems, with potential applications in real-time ride-hailing platform management, demand–supply balancing, and operational decision-making in the emerging autonomous taxi era.
However, one limitation of this study is that the evaluation is conducted on a relatively small city-scale network, where demand and supply are relatively evenly distributed, which may not fully reflect the complexity of larger metropolitan areas. Future work will extend the framework to larger cities with more heterogeneous demand and supply patterns to further assess scalability and robustness.

References

Alonso-González, M. J., Cats, O., Van Oort, N., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2021). What are the determinants of the willingness to share rides in pooled on-demand services? *Transportation*, *48*(4), 1733–1765.

Bahrami, S., Nourinejad, M., Nesheli, M. M., & Yin, Y. (2022). Optimal composition of solo and pool services for on-demand ride-hailing. *Transportation Research Part E: Logistics and Transportation Review*, *161*, 102680.

Chen, C., Yao, F., Mo, D., Zhu, J., & Chen, X. M. (2021). Spatial-temporal pricing for ride-sourcing platform with reinforcement learning. *Transportation Research Part C: Emerging Technologies*, *130*, 103272.

Ge, S., Zhou, X., & Qiu, T. (2024). 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.

Hu, J., Wellman, M. P., et al. (1998). Multiagent reinforcement learning: theoretical framework and an algorithm. In *Icml* (Vol. 98, pp. 242–250).

Huang, J., Huang, L., Liu, M., Li, H., Tan, Q., Ma, X., ... Huang, D.-S. (2022). Deep reinforcement learning-based trajectory pricing on ride-hailing platforms. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *13*(3), 1–19.

Jacob, J., & Roet-Green, R. (2021). Ride solo or pool: Designing price-service menus for a ride-sharing platform. *European Journal of Operational Research*, *295*(3), 1008–1024.

Kang, S., Mondal, A., Bhat, A. C., & Bhat, C. R. (2021). Pooled versus private ride-hailing: A joint revealed and stated preference analysis recognizing psycho-social factors. *Transportation Research Part C: Emerging Technologies*, *124*, 102906.

Lei, Z., & Ukkusuri, S. V. (2023). Scalable reinforcement learning approaches for dynamic pricing in ride-hailing systems. *Transportation Research Part B: Methodological*, *178*, 102848.

Li, Z. (2023). Leveraging explainable artificial intelligence and big trip data to understand factors influencing willingness to ridesharing. *Travel Behaviour and Society*, *31*, 284–294.

Liu, Y., Bansal, P., Daziano, R., & Samaranayake, S. (2019). A framework to integrate mode choice in the design of mobility-on-demand systems. *Transportation Research Part C: Emerging Technologies*, *105*, 648–665.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Shi, B., Cao, Z., & Luo, Y. (2022). A deep reinforcement learning based dynamic pricing algorithm in ride-hailing. In *International conference on database systems for advanced applications* (pp. 489–505).

Wu, W.-X., Sun, R., Liu, X.-M., Huang, H.-J., Tian, L.-J., & Shang, H.-Y. (2025). Optimal pricing and vehicle allocation in local ride-sharing markets with user heterogeneity. *Transportation Research Part C: Emerging Technologies*, *174*, 105084.

Zhang, K., & Nie, Y. M. (2021). To pool or not to pool: Equilibrium, pricing and regulation. *Transportation Research Part B: Methodological*, *151*, 59–90.

Zhang, Z., Yang, L., Yao, J., Ma, C., & Wang, J. (2024). Joint optimization of pricing, dispatching and repositioning in ride-hailing with multiple models interplayed reinforcement learning. *IEEE Transactions on Knowledge and Data Engineering*, *36*(12), 8593–8606.

Zhong, Y., Lan, Y., Chen, Z., & Yang, J. (2023). On-demand ride-hailing platforms with heterogeneous quality-sensitive customers: Dedicated system or pooling system? *Transportation Research Part B: Methodological*, *173*, 247–266.