A Top-to-Bottom Reposition Method for Ride-hailing Platforms

T. Chen[a], J. Liu[a], S. Feng[b,*] and J. Ke[a]

[a] Department of Civil Engineering, the University of Hong Kong, Hong Kong, China
ctj21@connect.hku.hk, jliu2049@hku.hk, kejintao@hku.hk
[b] Department of Logistics and Maritime Studies, the Hong Kong Polytechnic University, Hong Kong, China
sfengag@connect.ust.hk
* Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 14, 2024

Keywords: E-hailing; vehicle reposition; reinforcement learning; Monte Carlo Tree Search.

1 INTRODUCTION

Vehicle repositioning in the ride-hailing market addresses the significant spatiotemporal imbalance between supply and demand. Studies such as Xu *et al.* (2018) and Zong *et al.* (2018) highlight that a large portion of orders go unserved and drivers spend extensive periods without passengers. This has spurred research into developing effective repositioning algorithms (Zhu *et al.*, 2024, Chen *et al.*, 2024). Traditional grid-based methods, which direct drivers from one grid to another using the shortest route, often fail to optimize for critical metrics such as driver utilization or platform profit, focusing instead on minimizing travel time or distance. In this context, repositioning should aim primarily at securing the next passenger efficiently. Although Monte Carlo Tree Search (MCTS) has been effective in various applications, it struggles with temporal adaptability and demands extensive training time (Garg & Ranu, 2018). To address these challenges, we introduce a novel Top-to-Bottom Reposition Method (T2B-RM) integrating Reinforcement Learning (RL) and MCTS. The first stage leverages Mean-Field Multi-Agent RL (MF-MARL) to guide the regional movement of empty vehicles, deciding whether to stay or move to a new area. The second stage applies MCTS to determine the optimal routes within the targeted area, maximizing potential passenger pickups. This dual-layer approach ensures that vehicles are not only directed efficiently but also positioned optimally within target zones to enhance service availability. Our extensive experiments in Manhattan verify the efficacy of our method, demonstrating significant improvements in key performance metrics like platform total revenue (+2.4%) and order matching rate (+2.9%). These results substantiate our approach's potential in enhancing operational efficiencies in the ride-hailing industry.

2 METHODOLOGY

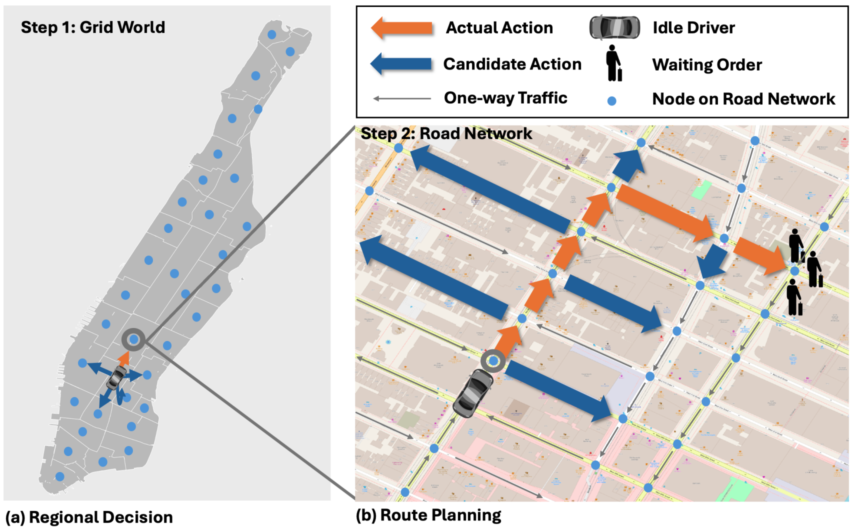


Figure 1 – *Illustration of Top-to-Bottom Reposition Framework*

We introduce the Top-to-Bottom Repositioning Method (T2B-RM) as depicted in Figure 1. This hierarchical repositioning strategy comprises two main stages:

1. **MF-MARL Stage**: In the initial stage, empty vehicle agents operate within a predefined grid-based environment. Each agent's state is characterized by $(\text{grid\_id}, \text{time\_step})$ and the possible actions include staying, or moving left, right, up, or down. The reward function is designed to compute the average revenue derived from drivers selecting identical actions. This stage leverages the Actor-Critic method of reinforcement learning to facilitate decision-making.

2. **MCTS Stage**: Following the selection of a target grid based on the RL policy, agents transition to the second stage, where MCTS is employed. This stage focuses on route generation, utilizing a heuristic to optimize the placement of orders in a timely manner. The implementation details of this stage are encapsulated in Algorithm 1.

Algorithm 1 MCTS

```
Initialize reward of each node i to 0, t ← 0
repeat
    Generate route 𝓡 via UCB policy
    if R finds a customer C then
        ∀i′ ∈ 𝓡\i, reward(i′) = order revenue - travelling cost
    else
        ∀i′ ∈ 𝓡\i, reward(i′) = - traveling cost
    end if
until maximum search depth is reached
Update reward X(i) for each node in 𝓡
t ← t + 1
```

To integrate the two stages, a weighted harmonic mean combining Q-value in RL stage, average Upper Confidence Boundary (UCB) in the target grid in MCTS stage (denoted as $\bar{w}$, calculated with Equation 1), and the change in $\bar{w}$ (denoted as Δw) is calculated and used to guide the RL agent (as shown in Equation 2).

$$UCB_i(t) = \overline{X}_i(t) + c\sqrt{\frac{\ln t}{n_i(t)}}, \quad \text{where } n_i(t) \text{ indexs \# of trials on node i at time t} \tag{1}$$

$$\pi^*(s) = \arg\max_{a \in A} \left(\beta_1 q_\pi(s,a) + \beta_2 \overline{w} + \beta_3 \Delta\omega\right), \quad \text{where } \beta_1 + \beta_2 + \beta_3 = 1 \tag{2}$$

3 RESULTS

3.1 Data

The study analyzes 65,955 ride-hailing orders from Manhattan, captured between 5:00 am and 10:00 am on May 4, 2015 (Monday). The focus is on the Manhattan borough's road network, which comprises 4,474 nodes and 9,682 edges. Orders with origins or destinations outside Manhattan were excluded. The dataset details are presented in Table 1.

Table 1 – *Dataset Statistics*

Dataset	# of Drivers	# of Orders	Order Sample Rate
Low Density	100	6,589	10%
High Density	500	32,967	50%

3.2 Model Results

In this study, we conducted comprehensive experiments on two datasets with different density levels, demonstrating that the T2B-RM consistently outperforms existing benchmarks by approximately 2% across all key performance metrics (as evidenced in Table 2). The MF-MARL model also showed enhanced performance over the standard MCTS model, which in turn exceeded the efficacy of a random strategy. This improvement was particularly notable in the low-density dataset.

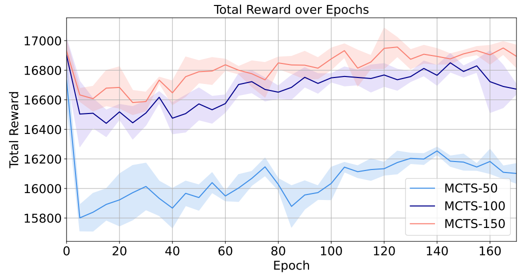
All models achieved similar outcomes in terms of order waiting times, with minimal variation in time allocation between matching and pickup in low-density scenarios.

Table 2 – *Experiment Results: On Low and High Density Dataset*

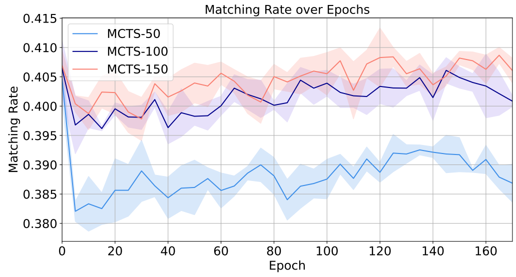
Models	Platform Revenue/$		Matching Rate		Occupancy Rate		Active Delivery Rate	
	Low	High	Low	High	Low	High	Low	High
Random	15,418	86,100	36.8%	41.16%	65.2%	69.03%	58.28%	64.99%
MCTS	16,852	90,574	40.51%	43.61%	71.95%	73.09%	63.79%	68.30%
MF-MARL	17,000	91,908	40.83%	43.92%	72.05%	74.2%	64.16%	69.45%
T2B-RM	**17,403**	**92,849**	**42.00%**	**44.5%**	**73.97%**	**74.53%**	**65.62%**	**70.16%**
	(+2.4%)	(+1.02%)	(+2.9%)	(+1.32%)	(+2.7%)	(+0.44%)	(+2.3%)	(+1.02%)

Models	Waiting Time/s		Matching Time/s		Pick-up Time/s	
	Low	High	Low	High	Low	High
Random	168	161	86	103	**82**	58
MCTS	165	162	**79**	101	86	61
MF-MARL	**164**	160	80	100	84	60
T2B-RM	**164**	**157**	**79**	**99**	85	**58**

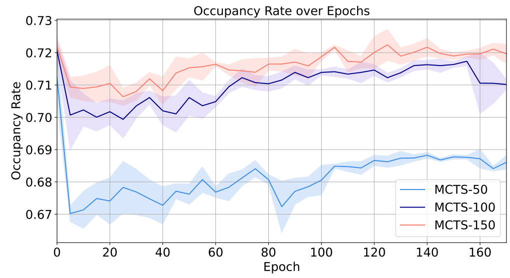
As shown in Figure 2, in the MCTS stage of our model, we tested various search depths and observed that larger search depths generally yield better results. However, the rate of improvement diminishes gradually with increased depth: there is a 6% improvement in platform revenue when increasing the depth from 50 to 100, but only around a 1% improvement when increasing it from 100 to 150.



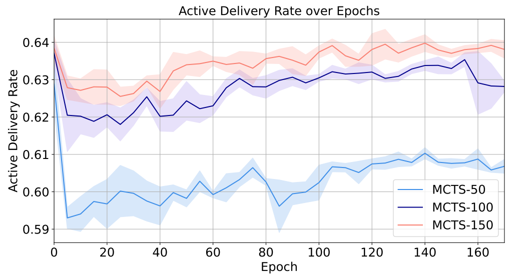
(a) *Total Reward*



(b) *Matching Rate*



(c) *Occupancy Rate*



(d) *Active Delivery Rate*

Figure 2 – *MCTS Maximum Search Depth: 50, 100 and 150*

4 DISCUSSION

In conclusion, our study addresses the crucial challenge of vehicle repositioning in the ride-hailing industry, which is vital for balancing supply and demand across urban areas. Previous methods that emphasize minimizing travel distances or times do not necessarily align with the primary goal of maximizing driver engagement with potential passengers. Our proposed T2B-RM innovatively combines an MF-MARL system with an MCTS system. This methodology not only guides vehicles to appropriate regions but also ensures optimal route selection within those regions to enhance passenger pickup rates. Our findings from extensive experiments conducted in Manhattan demonstrate that this approach can significantly improve key performance indicators such as platform revenue and matching rates. These results underscore the potential of our algorithm to transform operational strategies in the ride-hailing sector, making it a valuable tool for companies seeking to optimize their services and maximize efficiency.

References

Chen, Haoyang, Sun, Peiyan, Song, Qiyuan, Wang, Wanyuan, Wu, Weiwei, Zhang, Wencan, Gao, Guanyu, & Lyu, Yan. 2024. i-Rebalance: Personalized Vehicle Repositioning for Supply Demand Balance. *Proceedings of the AAAI Conference on Artificial Intelligence*, **38**(1), 46–54.

Garg, Nandani, & Ranu, Sayan. 2018. Route recommendations for idle taxi drivers: Find me the shortest route to a customer! *Pages 1425–1434 of: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.*

Xu, Ke, Sun, Luping, Liu, Jingchen, & Wang, Hansheng. 2018. An empirical investigation of taxi driver response behavior to ride-hailing requests: A spatio-temporal perspective. *PloS one*, **13**(6), e0198605.

Zhu, Pengbo, Sirmatel, Isik Ilber, Ferrari-Trecate, Giancarlo, & Geroliminis, Nikolas. 2024. A Coverage Control-Based Idle Vehicle Rebalancing Approach for Autonomous Mobility-on-Demand Systems. *IEEE Transactions on Control Systems Technology.*

Zong, Fang, Wu, Ting, & Jia, Hongfei. 2018. Taxi drivers' cruising patterns—Insights from taxi GPS traces. *IEEE Transactions on Intelligent Transportation Systems*, **20**(2), 571–582.