同济大学交通运输工程学院
COLLEGE OF TRANSPORTATION ENGINEERING
TONGJI UNIVERSITY

TRANSPORTATION RESEARCH BOARD

Author: Lingyun Zhong; 1953828@tongji.edu.cn
Author: Ye Li; jamesli@tongji.edu.cn
Corresponding Author:Meting Tu*; meitingtu@tongji.edu.cn

# Enhancing the Carbon Reduction Potential in Ridesplitting through a Reinforcement Learning: An case study of Chengdu

Paper ID: 24-03952

## INTRODUCTION

### Background

- The current level of sharing mobility in cities is relatively low and fails to meet expectations for energy conservation and emission reduction.
- Few studies focus on ride-splitting order matching that considers environmental benefits.
- Reinforcement learning can ensure long-term benefits of sequential decision-making, while the vehicle sharing network can effectively model ride-splitting order matching.

### Objective

- Combining reinforcement learning with the vehicle share-ability network model to solve ride-splitting order matching, achieving a balance between long-term platform benefits and environmental benefits

## METHODOLOGY

### Order Matching Framework

- Based on a vehicle share-ability network, identifying the possibility of ride-splitting between orders and the accessibility of drivers.
- Based on temporal-difference learning, updating the weights of the graph model.
- Based on the reduction principle, reducing the solution space and using the solver to obtain the optimal matching. The optimal matching result will be input into the value update formula of the temporal-difference learning algorithm to update our value function.

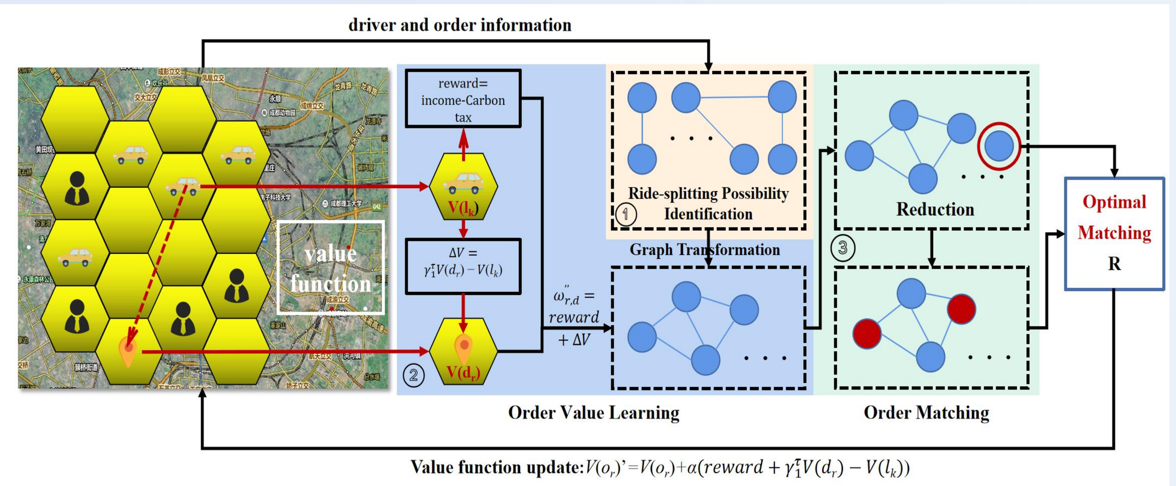


Figure 1 Order Matching Framework.

### Ride-splitting Identification Based on Vehicle Share-ability Network

Based on a vehicle share-ability network, identifying the possibility of ride-splitting between orders and the accessibility of drivers.

**Time window constraints(**After ride-splitting, the actual pick-up and delivery time should still be within the time window of the expected pick-up and delivery time and its maximum tolerated delay**)**:

$$t_i^s \le t_i^o \le t_i^s + \Delta$$
$$t_j^s \le t_j^o \le t_j^s + \Delta$$
$$t_i^e \le t_i^d \le t_i^e + \Delta$$
$$t_j^e \le t_j^d \le t_j^e + \Delta$$

$t_{i,j}^s$: *expected pick up time*; $t_{i,j}^o$: *actual pick up time*;
$t_{i,j}^e$: *expected delivery time*; $t_{i,j}^d$: *actual delivery time*;
$\Delta$: *maximum tolerence delay*

**Spatial path constraints**(The spatial constraints of the two orders for ride-splitting should satisfy that the general directions of the two orders should not have a large gap):

$$\theta_{ij} = arccos\frac{\overrightarrow{O_iD_i}\cdot\overrightarrow{O_jD_j}}{|\overrightarrow{O_iD_i}|\cdot|\overrightarrow{O_jD_j}|} \le [\theta]$$

$\theta_{ij}$: *the vector angle of the OD of two orders*;
$\overrightarrow{O_{i,j}D_{i,j}}$: *the OD vector*;
$[\theta]$: *the limitation of the vector angle of the OD of two orders*

**Vehicle share-ability network node weights calculation:**

$$\omega'_{r,d} = W - C_{tax}$$
$$C_{tax} = c_e \cdot Emission$$

$\omega'_{r,d}$: *vehicle share − ability network node wights*
$W$: *Calculated by DIDI*; $C_{tax}$: *total carbon tax;*
$c_e$: *the carbon tax rate*; *Emission*: *Calculated by CMEM model*

### Order value learning approach based on temporal difference learning

A temporal-difference learning approach to learn historical data and implement policy evaluation for order matching.

**Temporal-difference learning**:

$$V(\delta(l_k)) \leftarrow V(\delta(l_k)) + \alpha[reward + \gamma^{\tau}V(\delta(D_i)) - V(\delta(l_k))]$$

**Node update**:

$$\omega''_{r,d} \leftarrow \omega'_{r,d} + \mu_{r,d}$$

**Reward funtion**:

$$Reward(R_t) = \sum_{(r,d)\in R_t} \omega''_{r,d}$$

**Long term optimization Objective**:

$$max \sum_{t=1}^{t_{end}} Reward(R_t)$$

### Order matching method based on single and two vertex reduction

A solver combined with single and two vertex reduction methods is proposed to solve th order-matching problem.

**Short term optimization Objective**:

$$max \sum_{(r,d)\in R_t} \omega''_{r,d}$$

***Optimization methods***:

*Single Vertex Reduction:*

Given a vertex $V_i$ ∈the intersection graph $G'$ and its neighbors $N(V_i)$
*if* $\omega''(V_i) \ge \omega''(N(V_i))$, $R_{max}(G') = R_{max}(G' \setminus (N(V_i) \cup V_i) \cap \{V_i\}$

*Two Vertex Reduction:*

Given two vertices $V_i$ and $V_j$ ∈ the intersection graph G' and the concatenation of their neighbors $N(V_i) \cup N(V_i)$
if $\omega''(V_k) \le \omega''(N(V_k))$ for each vertex $V_k$, and $\omega''(V_i) + \omega''(V_j) \ge \omega''(N(V_i)) + \omega''(N(V_j)$, $R_{max}(G')= R_{max}(G' \setminus (N(V_i) \cup N(V_j) \cup \{V_i, V_j\}) \cap \{V_i, V_j\}$

*Nerworkx solver:* to input the reduced graph into python networkx to solve for the minimum weighted vertex cover, and then take the complementary set to obtain the optimal order matching.

## Baseline Matching Model

- Highest Order Value First(HOVF): The algorithm starts from the drivers' point of view, so the orders are ranked based on the order initiation time and the order price (i.e., the driver's income), and the orders are matched according to the first-come, first-served criterion with the highest order price as the standard
- Vehicle Share-ability Network(VSN): This algorithm considers ride-splitting, along with the economic and environmental benefits of ride-splitting, but does not consider the sequential decision-making properties of the order matching.
- Reinforcement Learning Based Bipartite Graph Order Matching(RL-BGOM): This model uses time-series differential learning to learn historical order data and form the value grid, but does not consider the ride-splitting behavior, while considering the economic and environmental benefits.

## RESULTS

### Results of Matching Time consumption

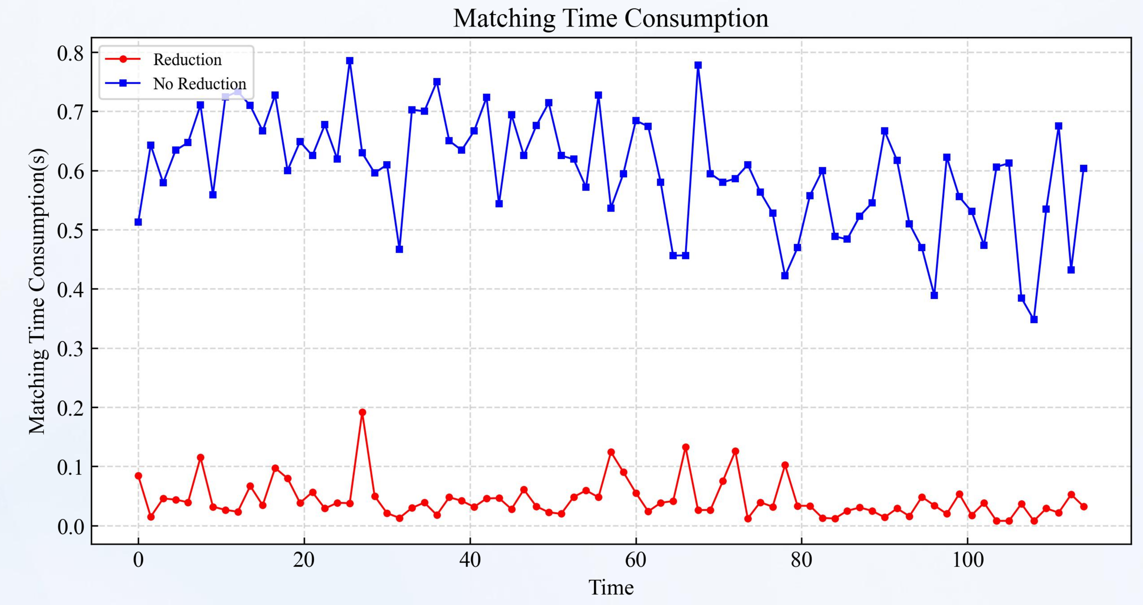


Figure 2 Matching Time Consumption

### Results of Net Revenue Compared to RL-BGOM

| Car Fleet | Net Revenue (Our Solution) | Net Revenue (RL-BGOM) | Net Revnue Increase(%) |
|---|---|---|---|
| 3000 | 9607.25 | 8890.83 | 8.05 |
| 4000 | 10626.14 | 10202.77 | 4.15 |
| 5000 | 10720.51 | 10296.64 | 4.12 |
| 6000 | 10830.48 | 10456.85 | 3.57 |
| 7000 | 10853.56 | 10481.96 | 3.55 |
| 8000 | 10974.53 | 10462.38 | 4.90 |

Table 1 Results of Net Revenue Compared to RL-BGOM

### Results of Carbon Emission Compared to RL-BGOM

| Car Fleet | Carbon Emission (Our Solution) | Carbon Emission (RL-BGOM) | Carbon Emission Increase(%) |
|---|---|---|---|
| 3000 | 12295.73 | 11733.69 | -4.79 |
| 4000 | 12262.26 | 15368.43 | 20.21 |
| 5000 | 12363.10 | 16912.45 | 26.90 |
| 6000 | 12502.55 | 17003.13 | 26.47 |
| 7000 | 12507.04 | 17147.55 | 27.06 |
| 8000 | 12621.24 | 17098.90 | 26.19 |

Table 2 Results of Carbon Emission Compared to RL-BGOM

### Results of Total Carbon Emission Reduction

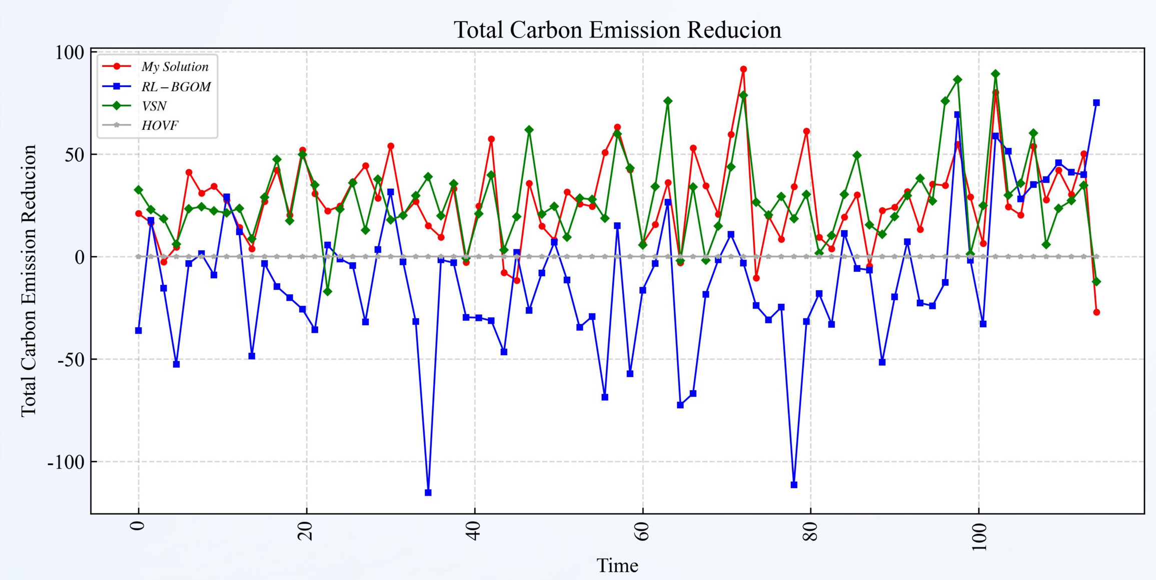


Figure 3 Results of Total Carbon Emission Reduction

### Results of Total Net Revenue

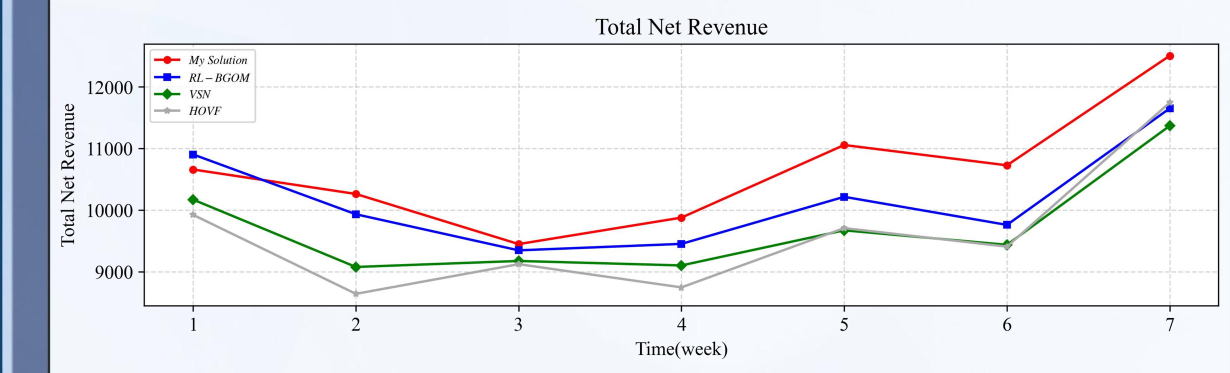


Figure 3 Results of Total Net Revenue

### Results of Total Peak Hour Carbon Emission Per Day in a Week

| DAY | My Solution | RL-BGOM | VSN | HOVF |
|---|---|---|---|---|
| Mon | 13102.6939 | 15275.0262 | 12571.0975 | 14913.9876 |
| Tue | 11525.6308 | 13998.5229 | 11300.5593 | 13391.8868 |
| Wed | 11814.0543 | 14039.9487 | 11434.9095 | 13876.2730 |
| Thu | 11599.4412 | 14080.7053 | 11520.3620 | 13741.3490 |
| Fri | 12374.3914 | 14726.0721 | 11984.0892 | 14614.6787 |
| Sat | 13025.2262 | 14380.2953 | 11664.1927 | 14204.4661 |
| Sun | 13975.1743 | 16085.0728 | 13684.6225 | 16878.7372 |

Table 3 Results of Total Peak Hour Carbon Emission

### Results of Order Completion Rate(%)

| Car Fleet | My Solution | RL-BGOM | VSN | HOVF |
|---|---|---|---|---|
| 3000 | 63.3% | 47.5% | 47.6% | 22.9% |
| 4000 | 63.1% | 52.3% | 54.5% | 28.1% |
| 5000 | 64.3% | 55.8% | 55.5% | 29.4% |
| 6000 | 65.1% | 57.3% | 55.7% | 29.4% |
| 7000 | 63.9% | 57.0% | 55.8% | 30.3% |
| 8000 | 65.1% | 57.3% | 56.0% | 31.2% |

Table 4 Results of Order Completion Rate

## CONCLUSIONS

- This paper combines reinforcement learning with the vehicle share-ability network model to model the ride-splitting order matching problem. At the same time, to solve the order matching problem, the reduction principle of single and two vertex is introduced to accelerate the algorithm.
- The method proposed in this paper can effectively model isolated matching batches as a sequential decision-making problem, ensuring long-term benefits.
- This article introduces environmental benefits into quantitative indicators, which can effectively promote energy conservation and emission reduction. At the same time, the results show that our method can achieve an effective balance between long-term benefits and environmental benefits.

## ACKNOWLEDGEMENT

This study was supported by the Fundamental Research Funds for the Central Universities [Grant numbers: 16002150138]; and the Science and Technology Commission of Shanghai Municipality [Grant number: 22dz1207500].