International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: xxxx

GPS DATA IN URBAN ONLINE CAR-HAILING: THE POTENTIAL REDUCTION OF INVALID EMISSION BY A STOCHASTIC OPTIMIZATION-BASED SMART MATCHING SYSTEM

Haoran Zhang[1*], Jinyu Chen[1], Wenjing Li[1], Qiong Zhang[2], Yin Long[2], Xuan Song[1], Ryosuke Shibasaki[1]

1 Center for Spatial Information Science, The University of Tokyo 5-1-5 Kashiwanoha, Kashiwa-Shi, Chiba 277-8568, Japan

2 Department of Environment Systems, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8563, Japan

ABSTRACT

Online ride-hailing is an emerging and popular transportation mode under the framework of sharing economy. Its emission issue has been focused by researchers for seeking a clean and efficient way to service passengers. In this paper, we based on stochastic optimization to proposed a smart matching method. The GPS dataset of Didi company in Chengdu city is employed to test the method. Compared with the shortest distance based matching method, the proposed method is able to reduce the invalid emission rate from 0.2834 to 0.2554 which proves the practicability of our method.

Keywords: GPS data, online car-hailing, invalid emission, stochastic optimization, matching system

NONMENCLATURE

Abbreviations	
COPERT	Computer Programme to calculate the Emissions from Road Transport
GPS	Global Positioning System
SO	Stochastic Optimization
Symbols	
Subscript and set	
$i \in I^t$	Driver agent set during the time-window t
$j \in J^t$	Existing and potential order set during the time-window t
J_r^t	Existing order set during the time-window t
J_p^t	Potential order set during the time-window t
$t \in T$	Time-window set
$s \in S$	Scenario set
Parameter	
$L_{i,j}^{t,s}$	Distance between the agent i and the order j during time-window t and at scenario s
τ^t	Time span of time-window t
wt_{max}	The maximum waiting time
M	A big positive value
Binary variable	
$A_{i,j}^t$	Matching variable for agent i and the order j during time-window t
Continuous variable	
wt_i^t	Accumulated waiting time of agent i at the beginning of time-window t

1. INTRODUCTION

Online ride-hailing is a new transportation mode under the framework of sharing economy. In China, online ride-hailing services now account for about 1% of the daily travel needs of the urban population, and there is plenty of room for further development. Effective development of online ride-hailing business can promote the utilization of traffic resources and urban traffic efficiency, improve the travel experience, and create more flexible employment opportunities for the society, so as to provide safe, convenient and comfortable travel services for residents. The matching system plays a key role in the online ride-hailing platform. A well-designed

Selection and peer-review under responsibility of the scientific committee of the 11th Int. Conf. on Applied Energy (ICAE2019).

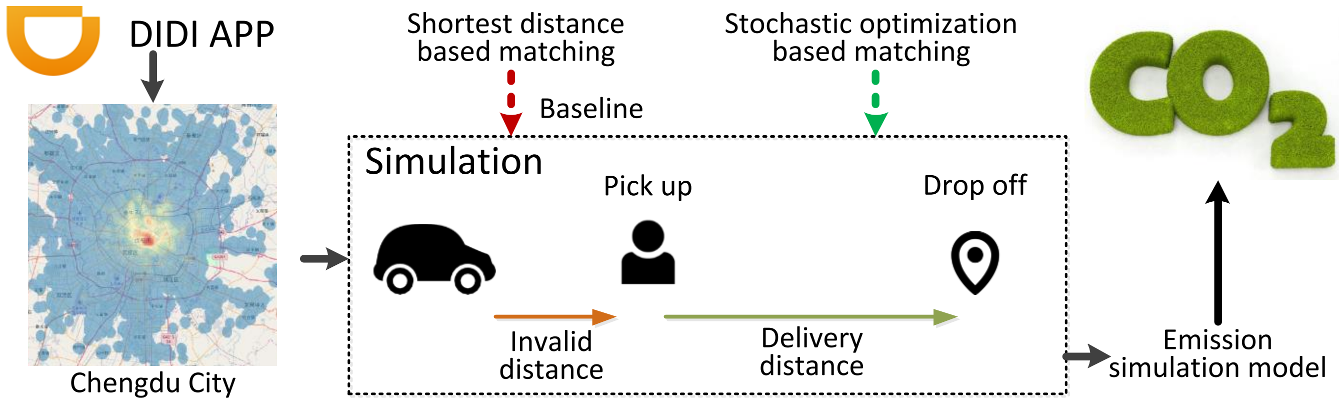

Fig 1 Framework of the model

matching system can not only improve the efficiency of the platform and reduce the traffic pollution and emission but also create a positive feedback to the passengers and drivers in user experience so as to promote the sustainable development of the online ride-hailing business. As the basis of the matching system development, the evaluation method to assess the efficiency of a theoretical matching system is the kernel of online ride-hailing research. However, large-scale and dynamic matching is a complex issue of system engineering. In addition, the online ride-hailing market is so complicated that it is hard to be simulated by a market model. These difficulties bring barriers to the researchers on developing the evaluation method and novel matching systems.

In this paper, we developed a stochastic optimization (SO)-based smart matching system and utilized the proposed evaluation method to quantificationally assess the potential reduction of invalid emission via this system in a real-world case.

2. RELATED WORKS

In the recent decade, with the development of sharing economy, the emerging shared transportation modes have become one of the hottest topics in the research fields of transportation and urban planning [1,2]. There are several researches focusing on developing the evaluation method and novel matching systems for online ride-hailing.

Based on the authoritative and heterogeneous data, Wu et al. [3] proposed an evaluation model of online car hailing services to evaluate impact on energy use and CO2 in China. The study mainly focused the macrocosmic data or information to assess the potential role of online ride-hailing in emission reduction. However, the effort of adopting a smart matching system on social and environmental improvement is hard to be quantificationally evaluated by these methods.

There are also several researchers forcing on developing the smart matching system to improve the efficiency of the online ride-hailing system. Based on reinforcement learning, Xu et al. [4] introduced a novel order dispatch algorithm in large-scale on-demand ride-hailing platforms. Through extensive offline experiments and online AB tests, the proposed approach delivers remarkable improvement on the platform's efficiency. However, the study did not a further work on the potential reduction of the emission by their proposed system.

Big data-driven analysis gives us a chance to uncover this issue. Sui et al. [5] utilized one month trajectory dataset and orders dataset, averagely covering around 7 thousand taxis with 0.3 million trips and 23 thousand Didi Chuxing Express vehicles with 0.1 million trips per day in Chengdu, China to make a comparative analysis on fuel consumption and emissions. A further work is to base on this emission simulation model to evaluate the impact of new matching methods on the improvement of system efficiency.

3. PROBLEM DESCRIPTION

3.1 Framework

The framework of this paper is as shown in Fig.1. The dataset we used is from the Didi Chuxing APP. The dataset contains both trajectory data of Didi Chuxing Express vehicles and order data of passengers. Here, we take Chengdu, China as the study case. These data then are inputted into the proposed evaluation model which contains three modules: simulation module, matching module, and emission module. After the computation, the potential reduction of invalid emission can be assessed.

3.2 Simulation module

We assume that every driver agent using the online ride-hailing system repeat a similar process. The process contains four actions. First, after the driver agent receives an order, the agent will drive the car from its parking place to the origin of the order. Since in this movement, the car is under the Void state and the driver does not earn any money, we define the distance of this movement as the invalid distance, and the corresponding caused emission as the invalid emission; After the agent arrive the origin of the order, the agent picks up the passenger; Then, the agent drive to the destination of the order. Since this movement is with the pay by the passenger, this distance is defined as delivery distance and the corresponding caused emission as the delivery emission; Finally, the agent drops off the passenger at the destination of the order. In order to avoid the excess invalid distance, we assume the agent will park the car at the destination for waiting for the next order and finish one process. Since the order information and the initial state of the driver agents can be obtained from the real-world case, with given a matching plan between driver agents and orders, the detailed movement of each agent can be simulated following the above process. In this paper, we focus on how to utilize the smart matching method to reduce the invalid emission.

3.3 Matching module

Here, we make a comparative analysis between two matching methods. The one is the shortest distance based matching which is the most classic and common used method. The online ride-hailing system will catch the state and the location of driver agents. When the system receives a new order from a passenger user, the system will automatically scan the drivers who are in the available state and find out the nearest driver and matching the order to the targeted driver.

The other one is the proposed method: SO-based matching method. In this method, we establish a stochastic optimization model which focuses on the probability of the location where the orders appearing in the future and can find out a global optimal matching plan for all driver agents and passenger users in a matching time window.

3.4 Emission module

In this paper, the Computer Programme to calculate the Emissions from Road Transport model (COPERT) is employed to evaluate the emission based on the GPS data and the simulation result. OPERT is a widely used emission model developed by European Environment Agency. Based on distinguishing vehicle categories, fuel types, road categories and other parameters, COPERT model determines the emissions.

3.5 Model requirement

Given: trajectory data of Didi Chuxing Express vehicles and order data of passengers.
Determine: matching plans (obtained via the two matching methods); simulated trajectories of each agent; average invalid distances of each process; average invalid emission of each process.
Assumptions: The travel demand of the passenger will not change; After dropping off a passenger, drivers will park their cars until being dispatched the next order and then drive directly to pick up passengers.

4. SO-BASED MATCHING METHOD

4.1 Stochastic optimization model

The target of this model is to minimize the total invalid distances of all driver agents. Therefore, the objective function of this model is shown as follow:

$$\min f^t = \sum_s \sum_i \sum_j A_{i,j}^t L_{i,j}^{t,s}$$
$$\forall i \in I^t, j \in J^t, t \in T, s \in S \quad (1)$$

Where, $J^t = J_r^t \cup J_p^t$.

The constraints of this model are shown as follow:

$$\textstyle\sum_i A_{i,j}^t = 1 \quad \forall i \in I^t, j \in J_r^t, t \in T \quad (2)$$
$$\textstyle\sum_j A_{i,j}^t = 1 \quad \forall i \in I^t, j \in J^t, t \in T \quad (3)$$
$$wt_i^t = \left(wt_i^{t-1} + \tau^t\right)\left(1 - \sum_j A_{i,j}^t\right)$$
$$\forall i \in I^t, j \in J_r^t, t \in T \quad (4)$$
$$wt_i^{t-1} - wt_{max} \le M \sum_j A_{i,j}^t$$
$$\forall i \in I^t, j \in J_r^t, t \in T \quad (5)$$

Constraint (2) represents that each existing order must be matched to one driver agent. Constraint (3) shows that each driver agent will be matched to an existing or a potential order. Constraint (4) calculates the accumulated waiting time of agent i. Constraint (5) indicates that when the accumulated waiting time is longer than a threshold value, the system must arrange a matching for this agent. The SO model can be solved by Monte Carlo method.

4.2 Matching algorithm

The matching algorithm is developed as follow:

Algorithm 1: Matching algorithm

Input the Didi Chuxing data;

Dividing the grid for Chengdu city;
Statistics on probability the potential order of each grid for each time-window;
Record the driver's initial position;
for t=1 to tn
do
- Based on the obtained probability to generate *n* sets of locations of potential orders of time-window *t;*
- $S = \{1, \cdots, n\}$;
- Calculate $L_{i,j}^{t,s}$ for each existing and potential orders;
- Obtain the matching plan via solving SO model;
- $n = n + 1$;

while (do not meet with the condition of convergence of Monte Carlo method)
- Do the simulation module;
- Do the emission module;

end
Output the result;

tn is the number of the last time-window of the simulation.

5. RESULT AND DISCUSSION

The massive GPS dataset used in this study is obtained from Didi Chuxing Cooperation. As the largest online ride-hailing company in China, Didi provide all kinds services including taxi hailing, private car hailing, and social ride-sharing for over 8.5 million people around Chengdu area by September 2016 (CBNData,2017), which has the highest market penetration. In this paper, the dataset is recorded from both registered taxis and regular Didi hailing cars as well as the passengers via a smartphone application as shown in Fig.2.


Fig 2 GPS data of online car-hailing

In this model, the duration of time-window is a sensitive parameter which will affect the result of optimization. Here a sensitivity analysis is made as shown in Fig.3. When the duration is set to 0, the SO based matching algorithm will cannot work and be transformed to the shortest distance based matching. From the result, we can see the proposed method is able to reduce the invalid emission rate from 0.2834 (shortest distance based matching) to 0.2554 (duration is 2.5 minutes).

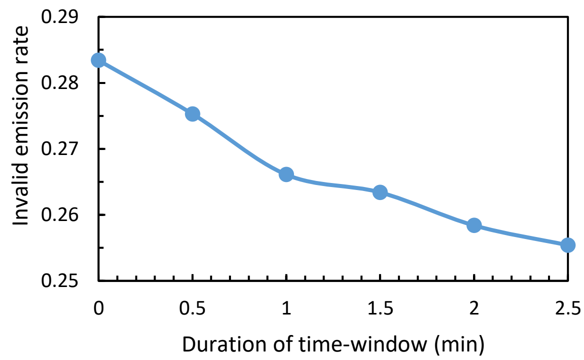

Fig 3 Sensitivity analysis

ACKNOWLEDGEMENT

This work was partially supported by the JST, Strategic International Collaborative Research Program (SICORP); Leading Initiative for Excellent Young Researchers (LEADER) Program, Grant in-Aid for Scientific Research B (17H01784) and Grant-in-Aid for Early-Career Scientists (19K15260) of the Japan Ministry of Education, Culture, Sports, Science, and Technology (MEXT).

REFERENCE

[1] Zhang, H., Song, X., Long, Y., et al. (2019). Mobile phone GPS data in urban bicycle-sharing: Layout optimization and emissions reduction analysis. Applied Energy, 242, 138-147.
[2] Yuan, M., Zhang, Q., Wang, B., et al. (2019). A mixed integer linear programming model for optimal planning of bicycle sharing systems: A case study in Beijing. Sustainable Cities and Society, 101515.
[3] Wu, T., Shen, Q., Xu, M., et al. (2018). Development and application of an energy use and CO2 emissions reduction evaluation model for China's online car hailing services. Energy, 154, 298-307.
[4] Xu, Z., Li, Z., Guan, Q., et al. (2018, July). Large-scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 905-913). ACM.
[5] Sui, Y., Zhang, H., Song, X., et al. (2019). GPS data in urban online ride-hailing: a comparative analysis on fuel consumption and emissions. *Journal of Cleaner Production*.