An integrated optimization methodology for minimizing operating costs of home delivery services in the O2O retail era

Abstract

In O2O (online-to-offline) retail, customers place orders online and the O2O retailer delivers products from stores to customers within the commitment time. The online orders usually fluctuate more sharply during peak periods: like lunchtime and dinner time, which may significantly affect the operating cost. Therefore, this paper transforms the blended workforce-staffing problem of O2O home delivery service into dynamic vehicle routing, fleet sizing, and workforce type assignment. And then, an integrated optimization model is proposed to minimize the delivery and employment cost of delivery services, which considers the dynamic features of online orders and heterogeneous drivers. Multi-period capacity vehicle routing problem with time window (MP-CVRP-TW), fleet sizing problem (FSP), and workforce type assignment problem (WTAP) are integrated into the model in a subtle way. To solve the integrated model, we develop an efficient hybrid algorithm, which blended the memetic algorithm, Hopcroft Karp algorithm, and branch-and-cut algorithm. The experiment results based on a leading Chinese O2O retailer indicate that our algorithm is more applicable for MP-CVRP-TW, which could obtain a high-quality solution with less time spent compared with Gurobi solver. An innovative constructing algorithm is proposed to generate directed acyclic graphs (DAG), and then Hopcroft Karp algorithm is applied to solve FSP. We also conduct a comprehensive simulation to present the improvements in total cost, with the different settings of commitment delivery time and workforce allocation strategy. The theoretical and numerical results would shed light on the delivery management of the O2O on-demand services.

Keywords: O2O retail; Dynamic vehicle routing planning; Fleet sizing; Heterogeneous drivers;

1 Introduction

A novel online-to-offline (O2O) retail mode is currently popular in China. Alibaba's Hema Fresh (Hema 2021), which was established in 2017 and is a pioneering example of the O2O retail mode in China, has more than 200 brick-and-mortar stores and offers more than 7,000 high-quality items to 25 million customers. The O2O retail mode implements both offline sales and online sales in an offline store. Online customers place orders and are offered instant delivery services, while traditional offline customers pick up the goods by themselves and check out at cashier desks (Chen, Fan et al. 2022). O2O instant delivery service is an important pillar to support the rapid growth of O2O retail mode. The number of instant delivery orders reaches 24.37 billion in 2020, and the average annual growth rate exceeds 30% (Dianwoda and iResearch 2019). Recently, the O2O retail mode has been quickly introduced by retailers.

However, the high cost of O2O instant delivery service is the bottleneck for the rapid development of O2O retail mode. The main components of costs in O2O instant delivery service are the transportation cost and employment cost (Dai and Liu 2020). Transportation cost make up the majority of the total logistics cost (Azi, Gendreau et al. 2012, Ulmer, Mattfeld et al. 2018, Zhang, Gu et al. 2021), which are mainly influenced by the route planning (Ritzinger, Puchinger et al. 2016, Peng, Zhang et al. 2020). With the increase in labor costs, employment cost is becoming a larger proportion of total logistics costs. It is necessary to consider the cost of employing delivery drivers in the actual operation. Therefore, how to minimize transportation and employment costs is a huge challenge for O2O retailers.

The problem is compounded by the dynamic features of the online order and the heterogeneous of workforces in the O2O retail. Firstly, online orders placed by customers fluctuate dramatically, with two distinct peak periods: lunch time and dinner time. If too much capacity is allocated to complete the orders during peak hours, there will be a lot of idle capacity during low hours (Dong and Ibrahim 2020). Secondly, there are three types of workforces that could be employed, which are in-house drivers, outsourcing drivers and crowdsourcing drivers. The different type drivers are heterogeneous in terms of service quality, delivery capability, and salary structure. To the best of our knowledge, few studies comprehensively analyzed the impact of the dynamic features of the online order and the heterogeneous of workforces on the cost of O2O instant delivery services (see the literature review in Section 2).

In addition, the commitment delivery time promised to customers by O2O instant delivery services is a key factor affecting the transportation cost and employment cost. In current practice, O2O retailers divide their business hours into segments at 30-minute intervals, and customers within a 3km radius of the retailer could select in which segments their orders would be delivered (Hema 2021). The 30 minutes is the delivery time promised to the customer by the retailer and is referred to as the commitment time. If the commitment delivery time is too long, it is not enough to attract customers, and if the commitment time is too short, it will spend a lot of logistics costs to ensure that the order is delivered on time. Different O2O retailers will set different commitment times, but how to trade-off the relationship between commitment time and logistics costs is ambiguous. To the best of our knowledge, few studies have delved into the relationship between commitment delivery time and logistics costs for the O2O retail.

Therefore, this paper puts forward an integrated optimization model to minimize the transportation and employment cost of O2O instant delivery services. In this methodology, an innovative integrated model is proposed, which integrated the model of the multi-period capacity vehicle routing problem with time window (MP-CVRP-TW), fleet sizing problem (FSP), and workforce type assignment problem (WTAP) in a subtle way to formulate the dynamic features of the online order and the heterogeneity of workforces in O2O retail. To solve the integrated model, an efficient hybrid algorithm is developed, which blended the memetic algorithm, Hopcroft Karp algorithm, and Branch-and-cut algorithm. In the case study, the experimental results validate the effectiveness of these algorithms. And the relationship between commitment delivery time and total cost of the O2O instant delivery services are revealed.

The contributions of this research are summarized as follows.

(1) We transform blended workforce-staffing problem of O2O instant delivery service into dynamic vehicle routing, fleet sizing, and workforce type assignment. Then construct an integrated model considering dynamic features of online orders and heterogeneous drivers to minimize transportation cost and employment cost of O2O instant delivery service. The integrated model neatly integrates multi-period capacity vehicle routing problem with time window (MP-CVRP-TW), fleet sizing problem (FSP), and workforce type assignment problem (WTAP).

(2) We put forward an efficient integrated algorithm incorporating memetic algorithm, Hopcroft Karp algorithm, and branch-and-cut algorithm, which could solve blended workforce-staffing problem effectively. Several efficient operators are proposed in memetic algorithm to solve MP-CVRP-TW, which could obtain the high-quality solution with less time spending comparing with Gurobi solver. An innovative constructing algorithm is proposed to generate directed acyclic graphs (DAG), and then Hopcroft Karp algorithm is applied to solve FSP.

(3) This paper draws the conclusion that employment costs represent a significant portion of total costs. O2O retailers should pay more attention to employment costs. Heterogeneous drivers could reduce employment costs effectively, but this would lead to a decline in service quality. Managers should trade-off cost and service quality when hiring heterogeneous drivers. More importantly, choosing the reasonable commitment delivery time could reduce total cost significantly. Gradually increasing commitment delivery time could reduce total cost significantly, but exceeding certain thresholds would be counterproductive and increase the risk of losing customers.

The remainder of this paper is organized as follows. Section 2 presents the related literature. Section 3 introduces the new methodology which includes the minimum path cover method and integer programming model, and hybrid algorithm including the Hopcroft-Karp algorithm and Branch-and-cut algorithm. Section 4 is the numerical study and sensitivity analysis. Section 5 provides some managerial implications for the O2O retailer. Section 6 concludes with final remarks and directions for future research.

2 Literature review

Individual studies on how to reduce the transportation cost and employment cost are productive nowadays, but not much have been achieved by integrating these issues together. Moreover, in the O2O retail, the dynamic features of the online order and the heterogeneity of workforces, making it more complicated to optimize the transportation cost and employment cost. In this paper, we propose an integrated optimization methodology to minimize the transportation and employment cost of O2O instant delivery services, which neatly integrate the model of the MP-CVRP-TW, FSP, and WTAP. The design of the integrated methodology draws great inspiration from dynamic vehicle route planning and fleet sizing, which are presented below.

As online orders are generated randomly in the O2O instant delivery services, dynamic vehicle route planning (DVRP) is an effective way to reduce transportation cost (Ojeda Rios, Xavier et al. 2021, Soeffker, Ulmer et al. 2022). Multi-period static programming is an effective strategy to address DVRP, which transformed DVRP into multiple static vehicle routing problems. In current practice, O2O retailers divide their business hours into segments at commitment-time intervals. For each commitment-time interval, this is a static capacity vehicle routing problem. Dai, Tao et al. (2019) provided a systematic method for the O2O platforms to optimize order assignment and routing. Chen, Fan et al. (2022) developed an online instant delivery scheduling model that considers the dynamic online order, decides the start time of deliveries and assigns riders to deliver online orders is proposed. And an online delivery scheduling algorithm is designed by introducing delivery rules with feasible delivery time windows. In this paper, we also apply the multi-period static programming to address DVRP.

For the fleet sizing problem (FSP), productive results from homogeneous fleet sizing issues. Vazifeh, Santi, Resta, Strogatz, and Ratti (2018) addressed the minimum fleet problem in on-demand urban, they provided a network-based solution to the following 'minimum fleet problem'. Santi et al. (2014) analyzed the optimal fleet sizing of the taxis in a city based on historical data. Green, Savin, and Savva (2013), and W.-Y. Wang and Gupta (2014) analyzed the nurse staffing problem, considering the random patient census rate and uncertain capacity due to nurses' absenteeism. Bidhandi, Patrick, Noghani, and Varshoei (2019) adopted the queueing network to analyze the capacity planning for a community care service network with a given capacity. Jing Dong (2020) applied the queuing model to analyze the cost-minimizing staffing decisions in service systems, including flexible agents and fixed agents. In addition, these researches focused on the fleet sizing problem are also based on homogeneous drivers (Azi, Gendreau, & Potvin, 2012; Voccia, Campbell, & Thomas, 2019; Yildiz & Savelsbergh, 2019). In this paper, we put forward a method according to Vazifeh et al. (2018) to solve the fleet sizing problem arising in O2O retailing.

Moghaddam, M., et al. (2020) proposed a model for solving the container drayage operations problem with heterogeneous fleet, multi-container sizes. In this paper, an integer programming model of the WATP is developed.

For the WTAP, the aim's to decide the number of heterogeneous workforces with minimal employment costs. Bhandari, Scheller-Wolf, and Harchol-Balter (2008) solved the constrained dynamic operator staffing problem that involves determining the number of permanent and temporary operators in a call center and provided an efficient algorithm to solve the problem. Pac, Alp, and Tan (2009) analyzed the integrated workforce capacity planning for the manufacturing sector. They proposed to hire contingent capacity to face the uncertainty and a certain number of workers to guarantee daily operations. Ata, Lee et al. (2016) address the volunteer staffing problem in a fresh food gleaning operation, which considers both uncertain food (demand) and labor supply. With the application of Internet technology in human resources, a new type of human resource, crowdsourcing resources, has been activated. Arslan, Agatz, Kroon, and Zuidwijk (2019) analyzed the parcel delivery service in a crowdsourced platform, considering full-time drivers and crowdsourced drivers. Yildiz and Savelsbergh (2019) employed queueing theory and approximation in instant meal delivery platforms to analyze the impact of service areas and service quality on capacity planning with crowdsourced couriers and company-employed drivers. Scherr, Neumann Saavedra, Hewitt, and Mattfeld (2019) proposed a service network design problem for the tactical planning of parcel delivery with autonomous vehicles. They considered a heterogeneous infrastructure where such vehicles may only drive in feasible zones but need to be guided elsewhere by manually operated vehicles in platoons. Zhao (2017) proposed a model for solving the location-routing problem with a heterogeneous fleet. In this paper, an integer programming model of the WATP is developed considering the heterogeneity of different types of drivers in terms of service quality, delivery capability, and salary structure.

Some studies integrating vehicle routing and fleet sizing have also given us great insight. Dai, Tao et al. (2019) provided a systematic method for the O2O platforms to optimize order assignment and routing. de Bittencourt, Seimetz Chagas et al. (2021) proposed a solution framework for the integrated problem of cargo assignment, fleet sizing, and delivery planning in logistics. Shehadeh, Wang et al. (2021) investigated the fleet sizing and allocation problem for the on-demand last-mile transportation systems. Dožić, Jelović et al. (2019) focused on the problem about the fleet sizing and fleet assignment in the case of single fleet (only one type of aircraft) and proposed a metaheuristic approach based on the variable neighborhood search methodology to solves both fleet sizing problem. Li, Li et al. (2019) analyze the order fulfillment problem for online retailers, which integrate order combination and vehicle routing problem. In this paper, an innovative graph-based approach is adopted to neatly integrate route planning and fleet sizing problems.

To the best of our knowledge, few studies has integrating these issues together to optimize the transportation cost and employment cost of O2O home delivery services. Moreover, existing optimization models are not applicable to this scenario due to the dynamic features of the online order and the heterogeneity of workforces in the O2O retail. In addition, the commitment delivery time promised to customers by O2O home delivery services is a key factor affecting the total cost, few studies have delved into the impacts of commitment delivery time on total logistics costs arising in O2O retail.

Therefore, this paper proposes an integrated optimization methodology to minimize the total cost of O2O home delivery services, which integrated the model of MP-CVRP-TW, FSP, and WTAP in a subtle way, and designed an efficient hybrid algorithm to solve the integrated model. The model and algorithm are provided in the following sections.

3 Model and algorithm

3.1 Problem description

The delivery process of a well-known Chinese O2O retailer is described here, which provides the basic foundation to construct the integrated model. This retailer has more than 200 offline stores in China and over 400,000 online orders per day. We took one of these stores as the subject of our study.

In current practice, the O2O retailer is available from 8:30 to 22:00. The business hours are divided into multiple segments with fixed time intervals. These online customers within a 3km radius of the retailer could choose a specific segment to receive their orders. The fixed time interval is the delivery time promised to the customer by the retailer and is referred to as the committed delivery time. Different O2O retailer has different committed delivery time, such as Costco in the U.S. promises same-day delivery, and Hema Fresh in China promises 30-minute delivery. To reduce transportation costs, these online orders within the committed delivery time are combined to generate several super orders by optimizing the vehicle routing plan. And then each super order is assigned to a driver. Drivers pick up goods at the stores and deliver them to their destination within the committed delivery time. Each driver should return to the stores at the end of the delivery. In O2O retail, three types of the workforce can be employed to fulfill orders, which are in-house drivers, outsourcing drivers, and crowdsourcing drivers. Different types of workforce have different service quality, delivery capability, and salary structure (Tao, Dai et al. 2020). Reasonable fleet sizing and workforce type allocation scheme could bring significant economic benefits to O2O retailers. In addition, motorcycles are the main transport tools for the O2O home delivery services. Motorcycles have limited loading capacity. For O2O retailer, they have strict standards for the quality of O2O home delivery service.

For the convenience of modeling, we make the following assumptions which are common in the literature in the O2O retail industry.

Assumption 1: For each type of driver, the average speed obtained from the historical orders is adopted to represent their delivery ability. If the driver's delivery speed is faster, then its delivery ability is stronger.

Assumption 2: For each type of driver, the average service score obtained from the customer review in historical orders is adopted to represent their service quality. If the customer's rating is higher, then its service quality is better

Assumption 3: There exists a requirement for the minimum average service score due to the customer requirement, platform strategy, and competition among platforms.

To make the models easier to understand, the mathematical notation and description in this paper are presented in Table 1.

Table 1 Notations

Set	Description
W	Set of the workforce's type, $W=\{1,2,\ldots r\}$.
G (V,E)	Set of vertex V and edge E in graph G, $V=\{v_1, v_2 \ldots v_n\}$, $E=\{e_1, e_2 \ldots e_m\}$.
P	Set of the minimum path cover, $P=\{p_1, \ldots p_i \ldots p_s\}$.
p_i	Set of the vertices that the i^{th} path covered, $p_i=\{v_j \ldots v_n\}$.
O	Set of orders in a typical day, $O=\{o_1, \ldots o_i \ldots, o_n\}$.
S	Set of super orders, $S=\{s_1, \ldots s_i \ldots, s_n\}$.
Parameters	**Description**
b_i	The number of orders in the p_i, $p_i \in P$.
r_{ik}	The kth vertex in the i^{th} path, $k \in M_i$, $i \in P$.
u_{ij}	Number of time-out orders when the p_i is assigned to j^{th} type of workforce, $i \in P$, $j \in W$.
q_j	Average service quality of the j^{th} type workforce, $j \in W$.
a_j	Average delivery ability of the j^{th} type workforce, $j \in W$.
t_i^g	Generation time of the i^{th} order; $i \in O$.

t_i^c	Completion time of the *i*th order; $i \in O$.
t_i^e	The earliest arrival time of the *i*th order; $i \in O$.
t_i^l	The latest arrival time of the *i*th order; $i \in O$.
l_i^p	The pick-up location of the *i*th order; $i \in O$.
l_i^d	The drop-off location of the *i*th order; $i \in O$.
m_i	The customer rating of the *i*th order; $i \in O$.
d_{ij}	Distance between the *i*th order and *j*th order, $i, j \in O$.
t_{ij}	Travel time between the *i*th order and *j*th order, $i, j \in O$.
fc_j	Fixed cost per day for the *j*th type workforce, $j \in W$.
oc_j	Payment per order for the *j*th type workforce, $j \in W$.
ct	Commitment time of home delivery service promised to the customer by the retailer.
bt	Business hours of the retail company.
β	The maximum capacity of the vehicle.
Θ	Transportation costs per kilometer
$\mathcal{R}$	Service time for driver delivery orders.
γ	The minimum number of paths for the DAG.
α	Minimum average service score requirements for the retail company.
ε	Minimum on-time order rate requirements for the retail company.

Variables	**Definition**
y_{ij}^k	If the *k*th route contains the *i*th order and the *j*th order, $y_{ij}^k = 1$; otherwise, 0. $i,j \in O$, $k \in S$.
v_i	Arriving time of the *i*th order, $i \in O$
x_{ij}	If the *i*th path is assigned to *j*th type of workforce, $x_{ij} = 1$;otherwise, 0. $i \in P$, $j \in M$.

3.2 Objective function

The aim of the integrated model is to minimize the transportation cost TC and employment cost EC as follows. In this case, we test one operational day cost of O2O home delivery services.

$$Min\ C = TC + EC$$

Based on the O2O business process, we propose the calculation method of *TC* for one operational day. O2O retailers divide their whole day's business hours *bt* into segments with the fixed commitment time *ct*, so the business hours are divided to *T* periods, where $T = bt/ct$. The whole day's online orders *O* are also categorized into multiple periods based on their time window, $O = \{O^1, \dots O^t \dots, O^T\}$. In each time period, the model of MP-CVRP-TW is applied to minimize the transportation cost TC^t. Therefore, the calculation formula of the one-day transportation cost is shown as follows. The detail calculation formula of TC^t in each period and *EC* will be introduced in the 3.3.1 and 3.3.3.

$$TC = \sum_{t \in T} TC^t$$

3.3 Integrated model

We transform blended workforce-staffing problem of O2O home delivery service into dynamic vehicle routing, fleet sizing, and workforce type assignment. Then construct an integrated model considering dynamic features of online orders and heterogeneous drivers to minimize transportation cost and employment cost of O2O home delivery service.

The integrated model neatly combines the model of MP-CVRP-TW, FSP, and WTAP by applying the graph theory. Specifically, the model of MP-CVRP-TW is developed to minimize the transportation cost. And then, we convert the FSP to a minimum path cover problem (MPCP) on the directed acyclic graphs (DAG) based on the graph theory. The novel DAG constructed method is proposed, which could adopt the output of MP-CVRP-TW to construct the DAG. It subtly combines the two models and make it possible to obtain the optimal fleet size for the whole day. Then, an integer programming model of WTAP is developed to determine the types of each drivers, which consider the heterogeneities of drivers in terms of service quality, delivery capability, and salary structure. Each sub-model of the integrated model is introduced as follows.

3.3.1 Model of MP-CVRP-TW

The model of MP-CVRP-TW is the sub-model of the integrated model, which is developed to minimize the transportation cost. The classical model of MP-CVRP-TW could be found in these literatures (Larrain, Coelho et al. 2019, Neves-Moreira, Amorim-Lopes et al. 2020). Considering the characteristics of the O2O retail, we modify the classical model of MP-CVRP-TW and show it below instead.

$$\text{Min}\, TC^t = \sum_{k\in S}\sum_{(i,j)\in O^t} d_{ij} y_{ij}^k * \Theta \tag{1}$$

s.t.

$$\sum_{(0,j)\in O^t} y_{0j}^k = 1 \qquad \forall k \in S \tag{2}$$

$$\sum_{(i,n+1)\in O^t} y_{i,n+1}^k = 1 \qquad \forall k \in S \tag{3}$$

$$\sum_{k\in S}\sum_{(i,j)\in O^t} y_{i,j}^k = 1 \qquad \forall i,j \in O^t \tag{4}$$

$$\sum_{(i,j)\in O^t} y_{i,j}^k \le \beta \qquad \forall k \in S \tag{5}$$

$$v_i + t_{ij} + \mathcal{R} - M(1 - y_{ij}) \leqslant v_j \qquad \forall i,j \in O^t \tag{6}$$

$$t_i^e \le v_i \le t_i^l \qquad \forall i \in O^t \tag{7}$$

$$\sum_{(i,j)\in O^t} y_{ij}^k - \sum_{(j,i)\in O^t} y_{ji}^k = 0 \qquad \forall k \in S \tag{8}$$

$$y_{ij}^k \in \{0,1\} \qquad \forall i,j \in O^t, k \in S. \tag{9}$$

$$v_i \ge 0 \qquad \forall i \in O^t \tag{10}$$

The objective function (1) aims to minimize the transportation cost for the t^{th} time period. Constraints (2) and (3) indicate that the vehicle starts and ends at the depot. Constraint (4) represents that every order must be served. Constraint (5) indicates the sum of the capacities of all orders in the super order could not exceed the capacity of the vehicle. Constraints (6) and (7) show the arriving time of the order must be within the time windows. Constraint (8) is the flow balance constraint. Constraints (9) and (10) denote the range of values of the decision variables.

For the t^{th} time period, the output of the model results in multiple routes, each covering multiple orders. In this paper, we refer to each route as a super order. After collecting the output of the model for each time period, we could obtain a complete set of super orders which contains the orders for the whole day. These super orders will be adopted as vertex to construct the directed acyclic graphs (DAG) for the model of FSP

3.3.2 Model of FSP

The model of FSP is the sub-model of the integrated model, which aims to find the minimum number of drivers required to fulfill the whole day's online orders. We transform the FSP problem into a minimum path cover problem according to the graph theory. The main reason is that the FSP is NP-hard, these accurate methods based on mathematical programming model can't handling large-scale cases (Vazifeh, Santi et al. 2018, Yu, Redi et al. 2020). Therefore, this paper employed the graph theory to solve FSP.

The minimum path cover problem is formally defined as follows. In a graph $G = (V, E)$, define a path p_i in G to be a sequence of vertices $p_i = \{v_i \dots v_k\}$ and the $P = \{p_1, \dots p_i \dots p_s\}$ to be a set of paths which can cover the graph G where the vertex is covered only once. The objective of the minimum path cover problem is to find a set P with minimum number of paths to cover all vertices. In this scenario, the minimum number of paths indicates the minimum number of drivers required to fulfill these orders. The vertices covered by each path are the orders that the driver needs to deliver.

Since the minimum path cover problem could be solved efficiently on directed acyclic graphs (DAG) (Boesch and Gimpel 1977), we propose a novel method to construct the DAG for the FSP in O2O retail . The traditional method for constructing a DAG is shown in Figure 1. In the general DAG, each order is represented as a vertex, and the arcs between these vertices are obtained according to specific criteria. However, the traditional method of constructing a DAG ignore the order combination characteristics of O2O home delivery services. Therefore, in our proposed method, the super orders obtained from the output of MP-CVRP-TW are used as vertices of the DAG. And the criterion for determining whether an arc exists is proposed (see the Hopcroft Karp algorithm in Section 3.4.2). The novel method establishes a connection to the rich applied mathematics and computer science field of graph algorithms.

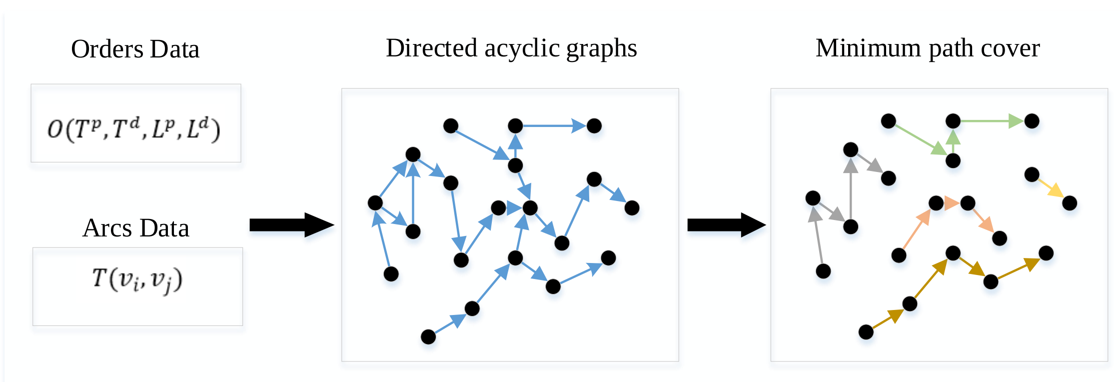


Figure 1 The traditional method for constructing a DAG

The output of the model is the minimum drivers required to fulfill these super orders, which be adopted to construct the labor resource constraint in the model of WTAP.

3.3.3 Model of WTAP

The model of WTAP is the sub-model of the integrated model, which aims to minimize the employment cost. In O2O retail mode, there are three different type workforces could be employed. The heterogeneity of three type drivers is reflected in working period, salary, service quality, and delivery ability, as shown in **Table 2**.

Table 2 The employee cost of different driver types

Workforce type	Working period	Wage components	Service quality	Delivery ability
In-house	Full time period	$fc_1 * \sum_{i \in P} x_{i1} + oc_1 * \sum_{i \in P} b_i * x_{i2}$	q_1	a_1
Outsourcing	Specified time period	$fc_2 * \sum_{i \in P} x_{i2} + oc_2 * \sum_{i \in P} b_i * x_{i2}$	q_2	a_2
Crowdsourcing	Random time period	$oc_3 * \sum_{i \in P} b_i * x_{i3}$	q_3	a_3

In-house and the outsourcing workforce have the fixed working hours, while the working hours of crowdsourcing workforce is flexible. Their wage components are shown in the third column, which are adopted to construct the objective function. The service qualities are represented in the fourth column, which are applied to construct constraints. The fifth column is the delivery ability. The delivery ability is represented by the average speed obtained from the historical orders, and this data is adopted to predict whether an order is overdue (Liu, Jiang et al. 2020, Liu, He et al. 2021). Based on the above analysis, the model of WTAP is shown as follow:

$$\text{Min } EC = fc_1 * \sum_{i \in P} x_{i1} + fc_2 * \sum_{i \in P} x_{i2} + oc_2 * \sum_{i \in P} b_i * x_{i2} + oc_3 * \sum_{i \in P} b_i * x_{i3} \tag{11}$$

s.t

$$\sum_{j \in W} x_{ij} = 1 \qquad \forall i \in P \tag{12}$$

$$\left(\sum_{j \in W} \sum_{i \in P} q_j * b_i * x_{ij}\right) / \sum_{i \in P} b_i \geq \alpha \tag{13}$$

$$\left(\sum_{i \in P} b_i - \sum_{j \in W} \sum_{i \in P} u_{ij} * x_{ij}\right) / \sum_{i \in P} b_i \geq \varepsilon \tag{14}$$

$$x_{ij} = \{0,1\} \tag{15}$$

The objective function (11) aims to minimize the employment cost. Constraints (12) indicates each path must be served. Constraints (13) represents the average service score should be no less than the minimum service score requirement α. Constraints (14) indicates the on-time rate of orders should be no less than the requirement on-time rate. Constraints (15) represents the value of decision variables.

3.4 Hybrid algorithm

The hybrid algorithm is devised to solve the integrated model, which blended the memetic algorithm, Hopcroft Karp algorithm, and Branch-and-cut algorithm. The framework of the Hybrid algorithm is provided in Fig.2. For each t^{th} time period, the function **Memetic_algorithm** (O^t) minimize the transportation cost and output the super orders s_i for the t^{th} time period. After the loop is end, the set of super orders S for the whole day are generated (see lines 1-4, Algorithm 1). The set of super orders S is inputted into the function **Construct_algorithm** (S) to generate the DAG, so the function **Hopcroft_algorithm**(G) could determine the minimum number of drivers required to satisfy the whole day's online orders. Then, the function **Branch_and_bound_algorithn**(P) determine the type of workforce of each driver. The optimal solution for dynamic vehicle route planning, driver size and the type of each driver could be obtained.

Algorithm 1 The hybrid algorithm

1: **for** $t \leftarrow 1$ to T **do**
2: $\quad s_i \leftarrow$ **Memetic_algorithm**(O^t)
3: $\quad S \leftarrow S \cup s_i$
4: **end for**
5: DAG ← **Construct_algorithm**(S)
6: $y \leftarrow$ **Hopcroft_algorithm**(DAG)
7: Solution ← **Branch_and_bound_algorithm**(y)
8: **return** Solution

Figure 2 the framework of the Hybrid algorithm

3.4.1 Memetic algorithm

The framework of the memetic algorithm could be summarized in four steps (Peng, Zhang et al. 2020). The first step is the initial population that aims to generate feasible individuals. The second step is the solution selection. The third step is the local search to improve the quality of the individual. The fourth step is the population updating. Each step is described in detail below.

1) Initial population

A two-stage procedure is used to compose a set of solutions. The first stage is inspired by the cluster-first-route-second method, which divides customers into groups using the sweep algorithm (Miller 1974). In the second stage, the TSP solver is applied to optimize the vehicle route for each group. And if the capacity and the time is not enough to serve the customer, a new vehicle is employed. The more feasible solution can be generated by rotate the polar coordinate of the sweep algorithm.

2) Solution selection

The tournament selection strategy (Zhang, Qin et al. 2020) is adopted to select the solution, which involves running several "tournaments" among individuals chosen at random from the population and selects the winner. The selected solution will be inputted to the local search operators to improve its quality by searching its neighborhood area.

3) Local search operators

Four operators are applied to improve the selected solution. The first operator is the swap operator, which exchanges the positions of any two customers to obtain a new solution. Then, the second one is the move operator, it moves the customer from its current position to another position. The third one is the two-opt operator. This operator exchanges the segment within two vehicle routes. Finally, the last one is the reverse operator, which used to reverses the order of the specific segment. To sum up, if the new solution is feasible and has the better quality, the operators are conducted.

4) Population updating

In each iteration, new solutions are generated by the local search operators. The new population is update by these new solutions. To maintain the diversity of the population, the probability of crossover and mutation operators are applied in the population.

3.4.2 Constructing algorithm to generate DAG

1) Determine the Vertex of DAG

The set of super order *S* are output by the memetic algorithm. Since these super orders are applied to construct the directed acyclic graphs (DAG), the information of each super order s_i (T_i^e, T_i^l, T_i^c) should be determined. The earliest arrival time T_i^e and the latest arrival time T_i^l are the time windows of the super order, which are equivalent to the time windows of the t^{th} period. The completion time is noted as T_i^c, which corresponds to the time when the driver returns to the warehouse.

2) Determine the arcs of DAG

If there is an edge between two vertices in the DAG, it means that one driver can fulfill these two super orders. A driver can serve only one super order at one time, and the driver has to deliver the next super order after fulfilling previous super order. We adopt the following criterion to determine if there is an edge between any two vertices, as shown in the formula (16). Only when the criterion holds, there will be an edge between two vertexes, and finally a DAG of the O2O home delivery service is formed.

$$T_i^c \leq T_{i+1}^e \tag{16}$$

3.4.3 Hopcroft Karp algorithm

After constructing the DAG, this paper applies the Hopcroft Karp algorithm to solve the minimum path cover problem. The problem of finding the minimum path cover on general graphs is NP-hard, but it can be solved efficiently on directed acyclic graphs (Boesch and Gimpel 1977) using the Hopcroft–Karp algorithm for bipartite matching (Hopcroft 1973). Hopcroft algorithm is a mature and efficient algorithm, the pseudo-code of the algorithm is shown in this article (Hopcroft 1973), so we not make redundant claims here.

3.4.4 Branch-and-cut algorithm

This paper adopts the Branch-and-cut algorithm to solve the model of WTAP. Branch-and-cut algorithm is an efficient method to solve the integer programming model. It's derived from the divide and conquer approach which break the problem into a series of smaller sub-problems, solve the smaller sub-problems, and then combine all local solutions of those sub-problems to form the optimal solution for the original problem.

In the Branch-and-cut algorithm, a typical way to represent such a divide and conquer approach is via an enumeration tree. Let $C = min\ \{cx\text{: } x \epsilon S\}$. Let $S = S_1 \cup \ldots \cup S_k$ be a decomposition of *S* into smaller sets, and let $C^k = min\{cx\text{: x}\epsilon S_k\}$ for *k=1,…,K*. then C = *min* c^k. However, complete enumeration branch of S_k is totally impossible for this problem. So the bound information could be applied to prune branches.

Let S= $S_1 \cup ... \cup S_k$ be a decomposition of S into smaller sets, and let $C^k = max\{cx\text{: x}\epsilon S_k\}$ for *k=1,…,K,* $\overline{C}^{\text{k}}$ be an upper bound on C^k and $\underline{C}^{\text{k}}$ be a lower bound on C^k. Then $\overline{C}$ = *max* $\overline{C}^{\text{k}}$ is an upper bound on C^k and $\underline{C}$ = max $\underline{C}^{\text{k}}$ is a lower bound on C^k.

Here we list the standard criterion of pruning branches as the following:

(1) Pruning by optimality: $C^t = \max\{cx\text{: x}\epsilon S_t\}$ has been solves.

(2) Pruning by bound: $\overline{C}^{\text{t}} \leqslant \underline{C}$.

(3) Pruning by infeasibility: $S_y = \emptyset$.

To decide on which variable should be branched, we calculate the relaxation optimal solution from the linear relaxation model of the original problem; and then select the variable which is closest to the integer to branch from the relaxation optimal solution. In addition, we adopt the depth-first strategies to search the node.

4 Case study

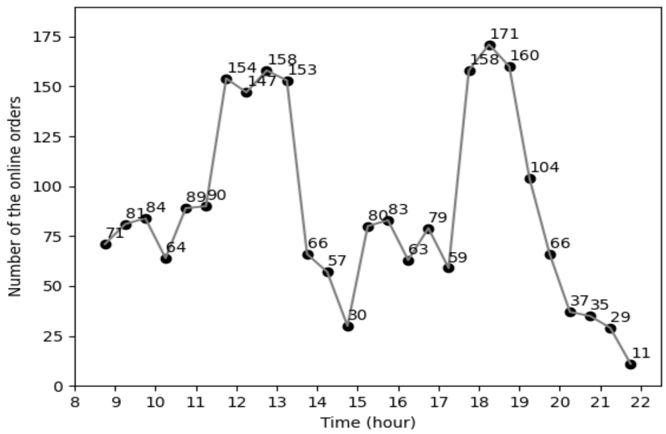
4.1 Data and parameters setting

In this research, we collect data from an offline store of a leading O2O retailer in China. The offline store locates in the Yubei district of Chongqing, China, from 8:30-22:00. We chose orders from a typical operational day as the experimental data to validate our approach. The experiment data contains 2380 online orders. Each order has these critical information $\{t_i^g, t_i^c, t_i^e, t_i^l, l_i^p, l_i^d, W, m_i\}$, as shown in **Table 3**.

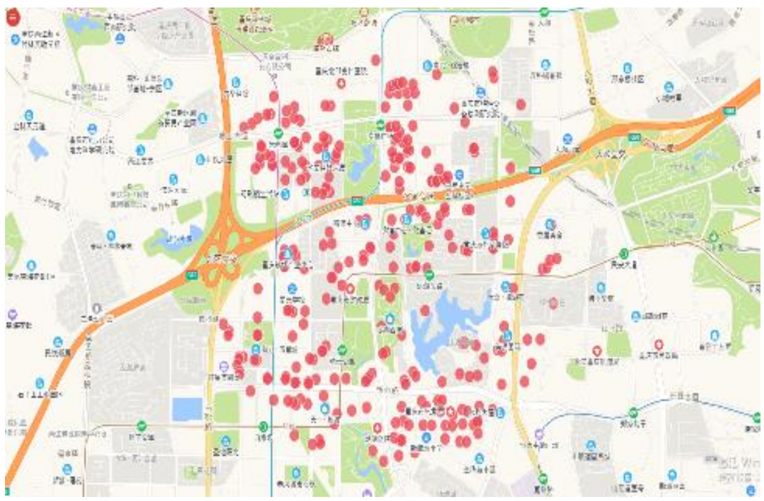
Table 3 The information of historical orders

Num	t_i^g	t_i^c	t_i^e	t_i^l	l_i^p	l_i^d	W	m_i
1	8:12	8:52	8:30	9:00	[106.539375,29.592201]	[106.531681, 29.576249]	1	100
2	8:08	8:56	8:30	9:00	[106.539375,29.592201]	[106.531811, 29.584772]	2	95
3	8:08	8:58	8:30	9:00	[106.539375,29.592201]	[106.532845, 29.589659]	1	100
…	…	…	…	…	…	…	…	…
2378	21:19	21:51	21:30	22:00	[106.539375,29.592201]	[106.533968, 29.589057]	3	80
2379	21:26	21:52	21:30	22:00	[106.539375,29.592201]	[106.537291, 29.589327]	1	95
2380	21:27	21:43	21:30	22:00	[106.539375,29.592201]	[106.534673, 29.580077]	1	100

The temporal and spatial distributions of these orders are shown in **Fig. 3**. It's obvious that the online orders fluctuate more sharply during peak periods: like lunchtime and dinner time. The demand for capacity is higher during peak periods, however, the demand for capacity during low periods is lower.



(a) Temporal distributions of orders



(b) Spatial distributions of orders

Fig.3 Temporal and Spatial distributions of orders

Besides, the parameter settings are shown in **Table 4**.

Table 4 Parameter settings

Parameters	Description	Unit	Value
W	Type of workforce.	--	[1,2,3]
q_1	Average service quality of in-house drivers.	Score	97
q_2	Average service quality of outsourcing drivers.	Score	93
q_3	Average service quality of crowdsourcing drivers.	Score	90
a_1	Average delivery ability of in-house drivers.	m/s	11
a_2	Average delivery ability of outsourcing drivers.	m/s	10
a_3	Average delivery ability of crowdsourcing drivers.	m/s	9
fc_1	Fixed cost per day for in-house drivers.	CNY	80

fc_2	Fixed cost per day for outsourcing drivers.	CNY	30
fc_3	Fixed cost per day for crowdsourcing drivers.	CNY	0
oc_1	Payment per order for in-house drivers.	CNY	1
oc_2	Payment per order for outsourcing drivers.	CNY	1.5
oc_3	Payment per order for crowdsourcing drivers.	CNY	2
β	The maximum capacity of the vehicle.	Orders	8
ct	Commitment time of the O2O retail	Minute	30
bt	Business hours of the retail company	Minute	[8:30-22:00]
Θ	Transportation costs per kilometer	CNY	0.4
$\mathcal{R}$	Service time for driver delivery orders	Minute	3
α	Minimum average service score requirements	Score	95
ε	Minimum rate of on-time orders	Percent	95%

4.2 Experimental Results

To verify the effectiveness of our proposed integrated optimization methodology, the hybrid algorithm is coded in Python 3.9, and run in Windows 10 with Intel Core i5-8250U CPU, with 1.80 GHz and 8.00 GB RAM. The Branch-and-cut algorithm is called from the Gurobi solver. The experiment results are provided in this section.

4.2.1 Minimal transportation costs

To evaluate the efficiency of the hybrid algorithm for solving the MP-CVRP-TW model, we also adopt the optimization solver Gurobi to solve the model. In Gurobi, the gap of the optimal solution is set to 5%, which could save the computation time. In addition, we adopt the hybrid algorithm to solve the MP-CVRP-TW model, recording the average results of ten calculations. These results are shown in Table 5.

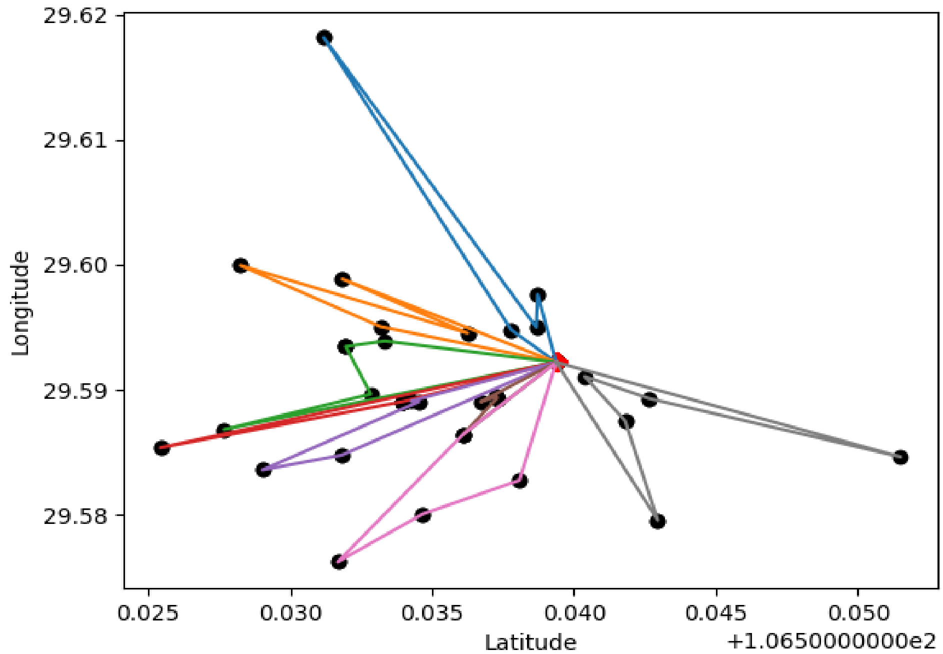
In Table 5, the first column represents the order of the time intervals. The second and third columns show the time windows of the time intervals. The fourth column represents the results of the Gurobi solver for solving the MP-CVRP-TW model, which includes the minimum transportation cost in CNY, the gap of the optimal solution, and the computation time in seconds. The fifth column records the results of the hybrid algorithm for solving the model, which also includes the minimum transportation cost, the gap of the optimal solution, and the computation time. The last two rows statistics the average and total dada of these indices.

According to the results, the hybrid algorithm is highly applicable to solve the MP-CVRP-TW model. For example, the average gap of Gurobi solver is 2.56%, but the average computation time is nearly 786 seconds. The computation time is too long for O2O home delivery service. Although the average gap of hybrid algorithm is 4.87%, the average computation time is only 89 seconds. The quality and the time consumption for generating the solution is more acceptable. Therefore, the hybrid algorithm is effective for solving the dynamical vehicle routing problem of O2O home delivery services. Furthermore, these results also confirm the validity of our proposed MP-CVRP-TW model.

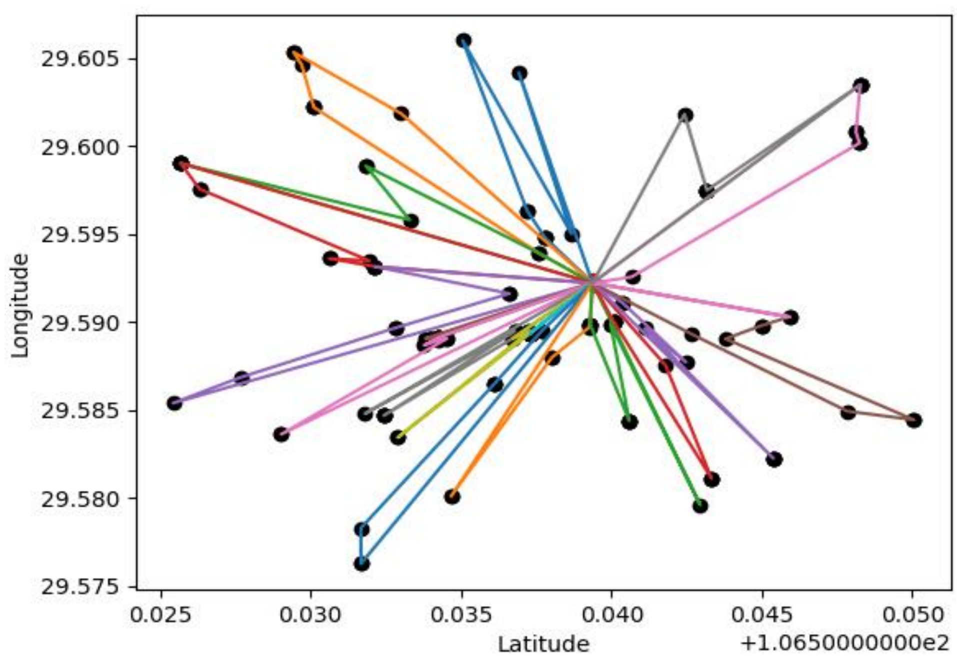
Table 5 Results of MP-CVRP-TW

I	t^e	t^l	Gurobi			Hybrid algorithm		
			Cost	Gap	Time	Cost	Gap	Time
1	8:30	9:00	11.07	0.50%	284	12.07	9.58%	88
2	9:00	9:30	11.46	2.80%	324	11.54	3.61%	89
3	9:30	10:00	13.57	4.30%	408	13.81	6.36%	95
4	10:00	10:30	8.90	0.80%	512	9.04	2.44%	73
5	10:30	11:00	12.90	3.00%	968	13.4	7.12%	100
6	11:00	11:30	13.15	2.70%	1080	13.17	2.92%	103
7	11:30	12:00	19.10	4.70%	1848	19.12	5.03%	154
8	12:00	12:30	16.85	3.30%	1352	17.23	5.73%	132
9	12:30	13:00	20.74	0.70%	1632	20.77	0.84%	169
10	13:00	13:30	20.33	2.90%	1224	20.71	4.90%	144
11	13:30	14:00	7.24	4.60%	528	7.73	11.85%	89
12	14:00	14:30	6.80	4.30%	912	6.88	5.74%	41
13	14:30	15:00	5.96	1.60%	360	6.03	2.24%	14
14	15:00	15:30	11.47	0.80%	780	12.04	5.79%	84
15	15:30	16:00	13.40	1.00%	332	13.49	1.69%	90
16	16:00	16:30	7.98	0.60%	504	8.01	0.83%	62
17	16:30	17:00	11.42	2.70%	632	11.53	3.75%	72
18	17:00	17:30	8.01	3.90%	708	8.05	4.64%	36
19	17:30	18:00	18.37	1.10%	1528	18.51	1.91%	156
20	18:00	18:30	22.57	4.40%	1368	22.95	6.35%	139
21	18:30	19:00	21.67	4.10%	1280	21.72	4.54%	136
22	19:00	19:30	15.00	2.50%	1016	15.13	3.49%	114
23	19:30	20:00	11.79	1.40%	792	11.85	1.96%	78
24	20:00	20:30	4.98	3.80%	296	5.12	6.96%	50
25	20:30	21:00	7.24	1.50%	280	7.35	3.04%	42
26	21:00	21:30	5.09	4.70%	232	5.2	7.16%	31
27	21:30	22:00	2.04	0.50%	44	2.26	11.12%	25
Average			--	2.56%	786	--	4.87%	89
Total			329.01	--	--	334.67	--	--

In each time interval, the super order is generated by the MP-CVRP-TW model, which would be applied to construct the DAG. After the calculation for each time interval was completed, 2380 orders were combined into 338 super orders. To show the super orders in different time intervals more visually, we select the super orders generated in 8:30-9:00 and 12:00-12:30, and present in Figure 4. In Figure 4, the point represents the location of the online order, and the cycle represents the route with the minimum distance to visit the orders. In other word, each cycle is a super order in Figure 4.



(a) Routing of super orders generated in 8:30-9:00



(b) Routing of super orders generated in 12:00-12:30

Figure 4 Routing of super orders generated in different time intervals

The necessary information of these super orders is obtained from the constructing algorithm, which is shown in **Table 6**. The first column represents the order of these super orders. The second and third column represent the time windows of each super order. The fourth column represent the complete time of the super orders. The fifth and sixth column represent the distance in km and the cost in CNY of each super order. The last column records the orders included in the super order. According to the information of super orders in the Table 6, the arc data of DAG could be obtained by the formula (16). Therefore, the DAG is constructed by the information of super orders and the arc data.

Table 6 The information of super orders

I	T^e	T^l	T^c	Distance	Cost	Orders
1	8:30	9:00	8:58	6.90	2.76	[28, 18, 15, 52, 55, 65]
2	8:30	9:00	8:56	1.53	0.61	[66, 16, 53, 9, 46, 6, 43, 5]
3	8:30	9:00	8:56	1.66	0.66	[7, 44, 8, 45, 13, 50, 51, 22]
…	…	…	…	…	…	…
335	21:00	21:30	21:15	2.57	1.03	[2361, 2347, 2366, 2352]
336	21:00	21:30	21:27	2.61	1.04	[2362, 2348, 2359, 2345, 2354, 2368, 2367, 2353]
337	21:30	22:00	21:42	2.51	1.00	[2370, 2369, 2375]
338	21:30	22:00	21:58	3.13	1.25	[2372, 2378, 2377, 2371, 2380, 2374, 2373, 2379]

4.2.2 Minimal fleet size

The hybrid algorithm is also applied to solve the FSP which are converted to a minimum path cover problem. The results of the minimum path cover problem are shown in **Table 7**. The first column represents the order of the path. The second shows the number of super orders of the path. The third column represents the number of orders included in the path. The fourth and fifth columns show the duration and the workload for delivering the orders in the path. The sixth column represents the total free time in the path. The last column shows the super orders included in the path.

To provide the minimum cover path more visually, the information of the path is shown in Figure 5. In Figure 5, each point represents a super order, and the number close to the point is the serial number of the super order. Each row indicates a path, and the super orders in the row are covered by the path in minute. Each column indicates the super orders generated in the specific time interval. The number of the minimum cover path is 22, it means that 22 drivers are needed to fulfil these orders.

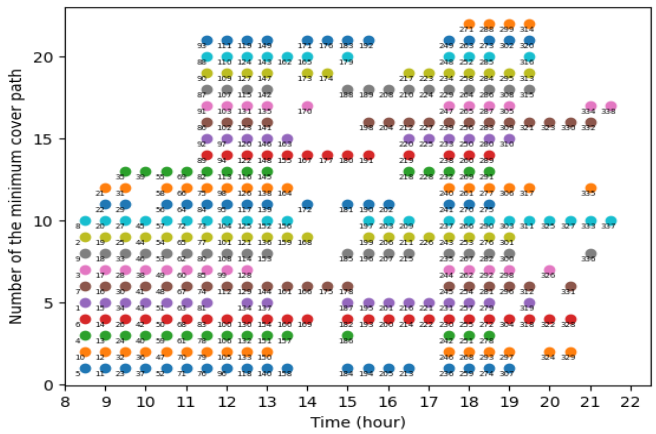


Figure 5 The minimum drivers needed to fulfil the whole day's online orders

Table 7 Results of the minimum path cover set

Path	Num_s	Num_o	Duration	Workload	Free time(min)	Super orders included in the path
1	19	140	8:30-19:30	660	90	[5, 11, 23, 37, 52, 71, 76, 96, 118, 140, 158, 184, 194, 205, 213, 236, 259, 274, 307]
2	16	109	8:30-21:00	750	270	[10, 12, 32, 36, 47, 70, 79, 105, 133, 150, 246, 268, 293, 297, 324, 329]
3	15	106	8:30-19:00	630	180	[4, 13, 24, 40, 59, 61, 78, 106, 132, 151, 157, 186, 242, 251, 278]
4	24	162	8:30-21:00	750	30	[6, 14, 26, 42, 50, 68, 83, 100, 130, 154, 160, 169, 182, 193, 200, 214, 222, 230, 255, 272, 304, 318, 322, 328]
5	18	131	8:30-20:00	690	150	[1, 15, 34, 43, 51, 63, 81, 134, 137, 187, 195, 201, 216, 221, 231, 257, 279, 319]
6	20	145	8:30-21:00	750	150	[7, 16, 30, 41, 48, 67, 74, 112, 129, 144, 161, 166, 175, 178, 245, 254, 281, 296, 312, 331]
7	14	109	8:30-20:30	720	300	[3, 17, 28, 38, 49, 60, 85, 99, 128, 244, 262, 292, 298, 326]
8	19	124	8:30-21:30	780	210	[9, 18, 33, 46, 53, 62, 80, 108, 114, 153, 185, 196, 207, 215, 235, 267, 282, 300, 336]
9	20	144	8:30-19:30	660	60	[2, 19, 25, 44, 54, 65, 77, 101, 121, 136, 159, 168, 199, 206, 211, 226, 243, 253, 276, 301]
10	23	153	8:30-22:00	810	120	[8, 20, 27, 45, 57, 72, 73, 104, 125, 152, 156, 197, 203, 209, 237, 266, 290, 303, 311, 325, 327, 333, 337]
11	15	108	9:00-19:00	600	150	[22, 29, 56, 64, 84, 95, 117, 139, 172, 181, 190, 202, 241, 270, 275]
12	15	111	9:00-21:30	750	300	[21, 31, 58, 66, 75, 98, 126, 138, 164, 240, 261, 277, 306, 317, 335]
13	13	82	9:30-19:00	570	180	[35, 39, 55, 69, 82, 113, 116, 145, 218, 228, 232, 269, 291]
14	13	103	11:30-19:00	450	60	[89, 94, 122, 148, 155, 167, 177, 180, 191, 219, 238, 260, 289]
15	11	72	11:30-19:30	480	150	[92, 97, 120, 146, 163, 220, 225, 233, 250, 280, 310]
16	16	107	11:30-21:30	600	120	[86, 102, 123, 141, 198, 204, 212, 227, 239, 256, 283, 309, 321, 323, 330, 332]
17	11	78	11:30-22:00	630	300	[91, 103, 131, 135, 170, 247, 265, 287, 305, 334, 338]
18	14	103	11:30-20:00	510	90	[87, 107, 115, 142, 188, 189, 208, 210, 224, 229, 264, 286, 308, 315]
19	13	92	11:30-20:00	510	120	[90, 109, 127, 147, 173, 174, 217, 223, 234, 258, 284, 295, 313]
20	11	79	11:30-20:00	510	180	[88, 110, 124, 143, 162, 165, 179, 248, 252, 285, 316]
21	13	90	11:30-20:00	510	120	[93, 111, 119, 149, 171, 176, 183, 192, 249, 263, 273, 302, 320]
22	4	29	18:00-20:00	120	0	[271, 288, 299, 314]

4.2.3 Minimal employment costs

In this case, we consider three labor allocation strategies arising in the O2O retail. The strategy 1 is that the O2O retailer just hire the in-house drivers. The strategy 2 is that the O2O retailer could hire the in-house drivers and outsourcing drivers. And the strategy 3 is that the three types of drivers are hired. For each strategy, the Branch-and-cut algorithm is applied to solve the WTAP which aims to minimize the employment cost.

The results are shown in Table 8. The first column represents the three labor allocation strategies. And the next three columns record the number of drivers in the optimal solution. The fifth column records the minimal employment cost in CNY. The sixth column represents the average score to fulfil these orders. The last column records the on-time rate of these orders.

Table 8 Results of the blended workforce-staffing

Strategy	In-house	Outsourcing	Crowdsourcing	E.Cost	A.Score	Rate
Strategy 1	22	-	-	4140	97	100%
Strategy 2	16	6	-	4069	96.4	98.5%
Strategy 3	17	4	1	4058	96.1	98.1%

In Table 8, no matter the O2O retailer choose which labor allocation strategy, the minimum number of drivers is 22. If more cheap type labors are available to delivery these online orders, its benefit to decrease the employment cost. Such as the strategy 3 could reduce the employment cost to 4058 CNY by hiring the 17 in-house drivers, 4 outsourcing drivers, and 1 crowdsourcing driver. However, the average score and on-time rate of the delivery service in the strategy 2 and 3 are lower than the strategy 1. The blended workforce-staffing scheme could cut the employment cost down, while compromising the quality of O2O instant delivery service.

4.3 Sensitivity analysis

Sensitivity analysis are presented in this section, which reveal the impact of commitment delivery time on transportation cost and employment cost under three different labor allocation strategies.

In this test, the commitment delivery times were set to [15, 20, 25, 30, 35, 40, 45] in minutes. And then, based on the completed time of the 2380 historical orders to divide them into the corresponding time intervals. As shown in Figure 6, the line graph shows the number of online orders in each time interval.

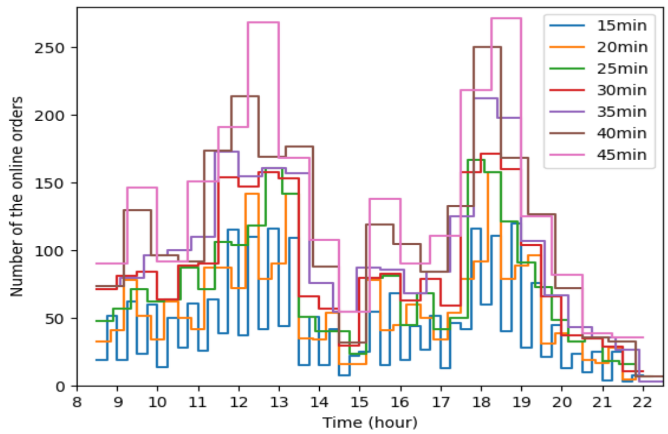


Figure 6 The distribution characteristics of online orders with different commitment delivery time

It is obvious that the commitment delivery time is larger, and the number of online orders fluctuates more sharply during the lunch and dinner time. When the commitment time is set to 45 minutes, online orders exceed 250 in the peak period and below 50 in the low period. However, when the commitment delivery time is set to 15 minutes, online orders is 120 in the peak period and below 50 in the low period.

4.3.1 The impact of the commitment delivery time on the transportation cost

To analyze the impact of the commitment delivery time on the transportation cost, the hybrid algorithm is applied to solve the MP-CVRP-TW model with different commitment delivery time. The experimental results are shown below.

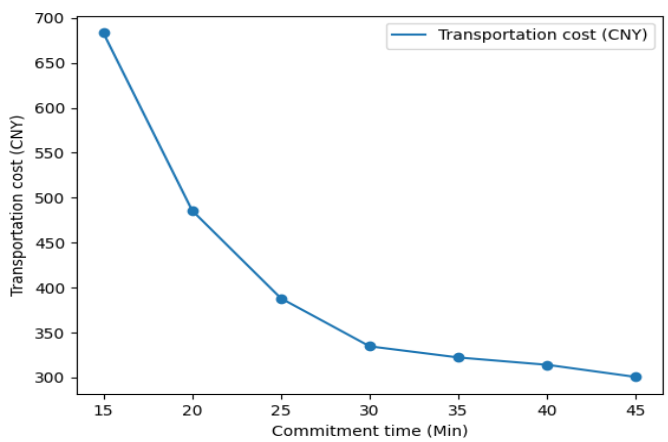


Figure 7 The impact of the commitment delivery time on the transportation cost

In Figure 7, we would draw the conclusion that the longer the commitment delivery time set by the O2O retailer, the lower of transportation cost. When the commitment delivery time is set to 45 minutes, the transportation cost is only about 300 CNY, which is just half of the transportation cost when the commitment delivery time is set to 15 minutes.

However, when we revisit the issue from the customer's perspective, customers' behaviors would change when setting a longer commitment delivery time. Therefore, increasing the commitment delivery time could reduce transportation costs, but there may be a risk of losing customers.

4.3.2 The impact of the commitment delivery time on the employment cost

To analyze the impact of the commitment delivery time on the employment cost, the hybrid algorithm is applied to solve the integrated model with different labor allocation strategies arising from the O2O retail. Strategy 1 is that the O2O retailer just hire the in-house drivers. Strategy 2 is that the O2O retailer could hire in-house drivers and outsourcing drivers. And strategy 3 is that the three types of workforce are available. The experimental results are shown below.

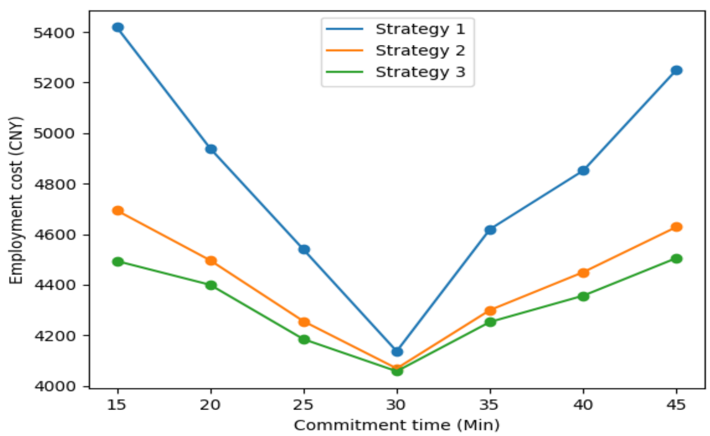


Figure 8 The impact of the commitment delivery time on the employment cost

In Figure 8, each line graph shows the impact of commitment delivery time on the employment cost with different labor allocation strategy. When the commitment delivery time gradually increases from 15 to 30 minutes, the employment cost decreases significantly. When the commitment delivery time exceeds 30 minutes, the employment cost increases dramatically. If the commitment delivery time is set relatively short, the total number of orders in each period will become smaller. However, the too-short commitment delivery time caused too few orders to be carried by each driver, the capacity of delivery tools would not be utilized sufficiently, and a large number of drivers would be needed to fulfill those orders. If the commitment delivery time is set too long, and long-time interval would accumulate more online orders. Although the capacity of delivery tools would be utilized sufficiently, a large number of drivers are still needed to fulfill these huge orders.

In this case, if the commitment delivery time is set to 30 minutes, which could obtain the minimum employment cost. In addition, cheaper workforce could effectively lower employment costs. Such as, the employment cost of strategy 3 would be reduced to nearly 4100 CNY.

4.3.2 The impact of commitment delivery time on total cost

The impact of the commitment delivery time on total cost, which is composed of transportation cost and employment cost, are investigated in this section. In Figure 9, it is obvious that the share of transportation cost is low in the total cost, and that transportation cost decrease with increasing commitment time. Employment cost is a major component of total cost approaching 90%.

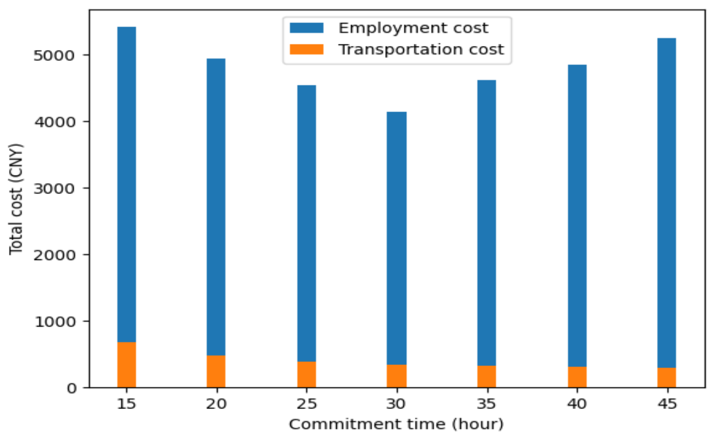


Figure 9 The impact of the commitment delivery time on the total cost with strategy 1

In Figure 10, each line graph represents the impact of commitment time on total cost with different labor allocation strategy. It is obvious that the labor allocation strategy has a significant impact on the total cost. It indicates that if O2O retailers have more types of cheaper drivers such as outsourcing drivers and crowdsourcing drivers, it would reduce labor costs significantly.

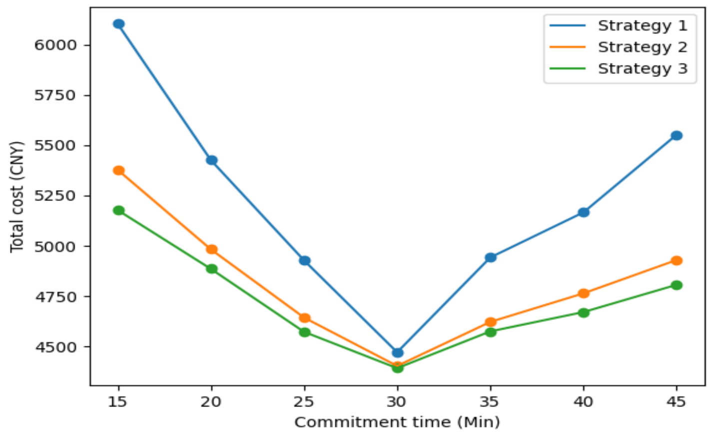


Figure 10 The impact of commitment delivery time on total cost with different labor allocation strategy

In addition, commitment delivery time has a significant impact on total cost. In this case, when commitment delivery time gradually increases from 15 to 30 minutes, total cost decreases significantly. When commitment delivery time exceeds 30 minutes, total cost increases dramatically. From Figure 10, we could draw the conclusion that the optimal commitment delivery time for the O2O retailer is 30 minutes, which has the minimum total cost with about 4500 CNY.

5 Conclusion

In this paper, we propose an integrated optimization methodology to minimize transportation costs and employment costs of HDS in the O2O retail industry. In this methodology, an innovative integrated model is proposed, which integrates the model of MP-CVRP-TW, FSP, and WTAP in a subtle way to consider the dynamic features of online orders and the heterogeneity of workforces in the O2O retail industry. And an efficient hybrid algorithm is developed. In the case study, the real data collected from a Chinese leading O2O retailer is adopted to verify the effectiveness of our methodology. The hybrid algorithm is more applicable for solving the dynamical vehicle routing problem arising in the O2O retail industry by comparing it with the Gurobi solver. And the experiment indicates that our methodology could efficiently reduce the operating costs of HDS.

More importantly, some valuable managerial implications are proposed in the sensitivity analysis.

(1) To minimize transportation costs, it is an effective measure to extend the committed delivery time. However, when the committed delivery time is extended beyond 30 minutes, the employment cost increases dramatically. How to set an appropriate committed delivery time is the key point.

(2) To minimize the employment costs, it is a valuable attempt to hire more types of drivers. It has to be stressed that the effectiveness of this method depends on the salary structure of different types of drivers. More types of cheap workforce could reduce employment costs effectively, such as outsourcing drivers and crowdsourcing drivers, but this would lead to a decline in service quality.

(3) To minimize the total costs, setting an appropriate committed delivery time is an effective measure. Based on our proposed method can help O2O retailers find the appropriate committed delivery time for their company. In this case, if the committed delivery time is set to 30 minutes, which could obtain the minimum operating costs of HDS with 4392 CNY.

Although this paper proposes an effective integrated optimization model to help O2O retailers minimize the total costs of HDS, there are still some limitations. More varieties of salary structures should be considered in future studies, which have a stronger impact on employment costs. Although the transportation cost is a relatively small part of the total cost, developing more efficient heuristic algorithms to obtain low-cost path planning solutions is still an effective way to reduce transportation costs. More heterogeneous features affecting the quality of delivery services should be considered, such as the difference in work shifts, rest breaks, and the length of work hours.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Reference

Abdulkader, M. M. S., Gajpal, Y., & ElMekkawy, T. Y. (2018). Vehicle routing problem in omni-channel retailing distribution systems. *International Journal of Production Economics, 196*, 43-55. doi:https://doi.org/10.1016/j.ijpe.2017.11.011

Ata, B., D. Lee and E. Sonmez (2016). "Dynamic staffing of volunteer gleaning operations." Available at SSRN 2873250.

Arslan, A. M., Agatz, N., Kroon, L., & Zuidwijk, R. (2019). Crowdsourced delivery—A dynamic pickup and delivery problem with ad hoc drivers. *Transportation Science, 53*(1), 222-235.

Azi, N., Gendreau, M., & Potvin, J.-Y. (2012). A dynamic vehicle routing problem with multiple delivery routes. *Annals of Operations Research, 199*(1), 103-112.

Bhandari, A., Scheller-Wolf, A., & Harchol-Balter, M. (2008). An exact and efficient algorithm for the constrained dynamic operator staffing problem for call centers. *Management Science, 54*(2), 339-353.

Bidhandi, H. M., Patrick, J., Noghani, P., & Varshoei, P. (2019). Capacity planning for a network of community health services. *European Journal of Operational Research, 275*(1), 266-279.

Boesch, F. T., & Gimpel, J. F. (1977). Covering Points of a Digraph with Point-Disjoint Paths and Its Application to Code Optimization. *Journal of the ACM, 24*, 192–198. doi:10.1145/322003.322005

Chang, S., Dong, Y., & Wang, X. (2021). Optimal shipping policy in retail competition and its effect on customers. *Electronic Commerce Research and Applications, 45*. doi:10.1016/j.elerap.2020.101020

Chen, J., Fan, T., Gu, Q., & Pan, F. (2022). Emerging technology-based online scheduling for instant delivery in the O2O retail era. *Electronic Commerce Research and Applications, 51*, 101115. doi:https://doi.org/10.1016/j.elerap.2021.101115

Cui, H., Chen, S., Chen, R., & Meng, Q. (2022). A two-stage hybrid heuristic solution for the container drayage problem with trailer reposition. *European Journal of Operational Research, 299*(2), 468-482. doi:https://doi.org/10.1016/j.ejor.2021.06.040

Dai, H., & Liu, P. (2020). Workforce planning for O2O delivery systems with crowdsourced drivers. *Annals of Operations Research, 291*(1), 219-245. doi:10.1007/s10479-019-03135-z

Dai, H., Tao, J., Jiang, H., & Chen, W. (2019). O2O on-demand delivery optimization with mixed driver forces. *IFAC-PapersOnLine, 52*(13), 391-396. doi:https://doi.org/10.1016/j.ifacol.2019.11.156

de Bittencourt, G. C., Seimetz Chagas, R. D., Silva, V. A., Peres Vianna, I. G., Longhi, R. P., Ribas, P. C., & Ferreira Filho, V. J. M. (2021). A solution framework for the integrated problem of cargo assignment, fleet sizing, and delivery planning in offshore logistics. *Computers & Industrial Engineering, 161*, 107653. doi:https://doi.org/10.1016/j.cie.2021.107653

Dožić, S., Jelović, A., Kalić, M., & Čangalović, M. (2019). Variable Neighborhood Search to solve an airline fleet sizing and fleet assignment problem. *Transportation Research Procedia, 37*, 258-265. doi:https://doi.org/10.1016/j.trpro.2018.12.191

Green, L. V., Savin, S., & Savva, N. (2013). "Nursevendor problem": Personnel staffing in the presence of endogenous absenteeism. *Management Science, 59*(10), 2237-2256.

Helsgaun, K. (2009). General k-opt submoves for the Lin–Kernighan TSP heuristic. *Mathematical Programming Computation, 1*(2-3), 119-163. doi:10.1007/s12532-009-0004-6

Hopcroft, J., & Karp, R. (1973). An n(5)/(2) algorithm for maximum matching in bipartite graphs. *SIAM J. Comput, 2*, 225–231

Jing Dong, R. I. (2020). Managing Supply in the On-Demand Economy: Flexible Workers, Full-Time Employees, or Both? *Operations Research, 68*(4), 1238-1264 doi:https://doi.org/10.1287/opre.2019.1916

Larrain, H., Coelho, L. C., Archetti, C., & Speranza, M. G. (2019). Exact solution methods for the multi-period vehicle routing problem with due dates. *Computers & Operations Research, 110*, 148-158. doi:https://doi.org/10.1016/j.cor.2019.05.026

Lei, C., Zhao, N., Ye, S., & Wu, X. (2020). Memetic algorithm for solving flexible flow-shop scheduling problems with dynamic transport waiting times. *Computers & Industrial Engineering, 139*, 105984. doi:https://doi.org/10.1016/j.cie.2019.07.041

Li, X., Li, J., Aneja, Y., Guo, Z., & Tian, P. (2019). Integrated order allocation and order routing problem for e-order fulfillment. *IISE Transactions, 51*(10), 1128-1150.

Liu, S., He, L., & Max Shen, Z.-J. (2021). On-Time Last-Mile Delivery: Order Assignment with Travel-Time Predictors. *Management Science, 67*(7), 4095-4119. doi:10.1287/mnsc.2020.3741

Liu, S., Jiang, H., Chen, S., Ye, J., He, R., & Sun, Z. (2020). Integrating Dijkstra's algorithm into deep inverse reinforcement learning for food delivery route planning. *Transportation Research Part E: Logistics and Transportation Review, 142*. doi:10.1016/j.tre.2020.102070

Miller, B. E. G. a. L. (1974). A heuristic algorithm for the vehicledispatch problem. *Oper. Res., 22*(340–349).

Ojeda Rios, B. H., Xavier, E. C., Miyazawa, F. K., Amorim, P., Curcio, E., & Santos, M. J. (2021). Recent dynamic vehicle routing problems: A survey. *Computers & Industrial Engineering, 160*, 107604. doi:https://doi.org/10.1016/j.cie.2021.107604

Pac, M. F., Alp, O., & Tan, T. (2009). Integrated workforce capacity and inventory management under labour supply uncertainty. *International Journal of Production Research, 47*(15), 4281-4304.

Peng, B., Zhang, Y., Lü, Z., Cheng, T. C. E., & Glover, F. (2020). A learning-based memetic algorithm for the multiple vehicle pickup and delivery problem with LIFO loading. *Computers & Industrial Engineering, 142*, 106241. doi:https://doi.org/10.1016/j.cie.2019.106241

Santi, P., Resta, G., Szell, M., Sobolevsky, S., Strogatz, S. H., & Ratti, C. (2014). Quantifying the benefits of vehicle pooling with shareability networks. *Proceedings of the National Academy of Sciences, 111*(37), 13290-13294.

Scherr, Y. O., Neumann Saavedra, B. A., Hewitt, M., & Mattfeld, D. C. (2019). Service network design with mixed autonomous fleets. *Transportation Research Part E: Logistics and Transportation Review, 124*, 40-55. doi:https://doi.org/10.1016/j.tre.2019.02.001

Shehadeh, K. S., Wang, H., & Zhang, P. (2021). Fleet sizing and allocation for on-demand last-mile transportation systems. *Transportation Research Part C: Emerging Technologies, 132*, 103387. doi:https://doi.org/10.1016/j.trc.2021.103387

Soeffker, N., Ulmer, M. W., & Mattfeld, D. C. (2022). Stochastic dynamic vehicle routing in the light of prescriptive analytics: A review. *European Journal of Operational Research, 298*(3), 801-820. doi:https://doi.org/10.1016/j.ejor.2021.07.014

Song, J., Yin, Y., & Huang, Y. (2017). A coordination mechanism for optimizing the contingent-free shipping threshold in online retailing. *Electronic Commerce Research and Applications, 26*, 73-80. doi:10.1016/j.elerap.2017.10.001

Taillard, É. D., & Helsgaun, K. (2019). POPMUSIC for the travelling salesman problem. *European Journal of Operational Research, 272*(2), 420-429. doi:https://doi.org/10.1016/j.ejor.2018.06.039

Vazifeh, M. M., Santi, P., Resta, G., Strogatz, S. H., & Ratti, C. (2018). Addressing the minimum fleet problem in on-demand urban mobility. *Nature, 557*(7706), 534-538. doi:10.1038/s41586-018-0095-1

Voccia, S. A., Campbell, A. M., & Thomas, B. W. (2019). The same-day delivery problem for online purchases. *Transportation Science, 53*(1), 167-184.

Wang, W.-Y., & Gupta, D. (2014). Nurse absenteeism and staffing strategies for hospital inpatient units. *Manufacturing & Service Operations Management, 16*(3), 439-454.

Wang, Y., Meng, Q., & Jia, P. (2019). Optimal port call adjustment for liner container shipping routes. *Transportation Research Part B: Methodological, 128*, 107-128. doi:https://doi.org/10.1016/j.trb.2019.07.015

Xu, N., Bai, S.-z., & Wan, X. (2017). Adding pay-on-delivery to pay-to-order: The value of two payment schemes to online sellers. *Electronic Commerce Research and Applications, 21*, 27-37. doi:10.1016/j.elerap.2016.12.001

Yildiz, B., & Savelsbergh, M. (2019). Service and capacity planning in crowd-sourced delivery. *Transportation Research Part C: Emerging Technologies*, 177-199.

Zhang, Z., Qin, H., & Li, Y. (2020). Multi-Objective Optimization for the Vehicle Routing Problem With Outsourcing and Profit Balancing. *IEEE Transactions on Intelligent Transportation Systems, 21*(5), 1987-2001. doi:10.1109/TITS.2019.2910274

Zhao, Q., Wang, W., & De Souza, R. (2017). A heterogeneous fleet two-echelon capacitated location-routing model for joint delivery arising in city logistics. . *International Journal of Production Research,, 56*(2), 1–19. doi:doi:10.1080/00207543.2017.1401235