Scheduling heterogeneous delivery tasks on a mixed logistics platform

Published Version:
Zhen, L.u., Baldacci, R., Tan, Z., Wang, S., Lyu, J. (2022). Scheduling heterogeneous delivery tasks on a mixed logistics platform. EUROPEAN JOURNAL OF OPERATIONAL RESEARCH, 298(2), 680-698 [10.1016/j.ejor.2021.06.057].

Availability:
This version is available at: https://hdl.handle.net/11585/849285 since: 2023-06-12

Published:

DOI: http://doi.org/10.1016/j.ejor.2021.06.057

Scheduling heterogeneous delivery tasks on a mixed logistics platform

Lu Zhen[1], Roberto Baldacci[2], Junyan Lyu[1], Shuaian Wang[3*], Zheyi Tan[1]

[1] *School of Management, Shanghai University, Shanghai, China*
[2] *Department of Electrical, Electronic, and Information Engineering, University of Bologna, Cesena, Italy*
[3] *Department of Logistics & Maritime Studies, The Hong Kong Polytechnic University, Hong Kong*

Abstract: Large e-commerce retailers usually establish their own logistics systems. Such systems make use of their own dedicated fleets but will also use a crowdsourced delivery mode by hiring occasional fleets. These mixed logistics systems with both dedicated and occasional fleets serve both retailers' internal delivery tasks and external tasks requested by local businesses. This paper studies the problem of scheduling heterogeneous (*internal* and *external*) delivery tasks on a mixed logistics platform with multiple depots and two types of vehicle (*dedicated* and *occasional*). A delivery task is executed by either a dedicated vehicle or an occasional vehicle. The dedicated vehicles depart from and return to the platform's depots; the occasional vehicles depart from their original location and pick up goods from depots or external pickup locations, fulfill the delivery tasks, and finish their route at the final delivery location. We propose mixed integer programming models and column generation-based solution methods to solve the problem. A computational study is conducted based on a series of randomly generated instances and real-world instances involving 20 depots, 200 internal customers, 40 external delivery tasks, and 70 dedicated and occasional vehicles. The results obtained demonstrate the efficiency of the column generation-based solution methods. Moreover, the effectiveness of the proposed models is validated by a significant cost saving in comparison to intuitive decision rules. A sensitivity analysis is also conducted to derive a number of managerial implications.

Keywords: OR in service industries; crowdsourced delivery; close-open mixed multi-depot vehicle routing problem; heterogeneous fleet; e-commerce logistics.

1. Introduction

The boom in retail e-commerce presents online retailers with new challenges in delivering a large number of time-critical orders in a short period. For example, in 2019, a total of 63 billion parcels were delivered to online shopping customers in China. On November 11 (China's annual "double-eleven" online shopping festival), the revenue generated by online shopping was 58.9 billion USD, from 1.66 billion orders (parcels). Aside from large volume, delivery time requirements are very strict. Some e-commerce retailers in China promise that if a customer makes an order before 11 pm, she/he will receive it before 3 pm the following day. Fulfilling a huge number of delivery tasks in a timely manner places a great deal of pressure on e-commerce retailers.

[*] Corresponding author. Email: wangshuaian@gmail.com (S. Wang)

Large e-commerce retailers in China such as JD.com and TMall.com have established their own logistics companies, such as JD Logistics and CAINIAO Logistics. To cope with operations challenges, e-commerce logistics companies are seeking ways to improve their capacity and efficiency, such as investing in new delivery technologies and adopting data mining tools to improve facility location decisions. Crowdsourced delivery has become a popular, cost-effective way for e-commerce logistics companies to improve their efficiency (Vazifeh et al., 2018; Alnaggar et al., 2019; Bergmann et al., 2020; Yang et al., 2020). Crowdsourced delivery takes advantage of the excess capacity of private vehicles to support delivery tasks. It may be possible to achieve faster and cheaper deliveries using this mode than with the traditional delivery mode (Gansterer and Hartl, 2018).

Large e-commerce logistics companies maintain dedicated fleets, and also adopt crowdsourced delivery to hire occasional fleets. For instance, about 10% of JD Logistics' fleet is fulltime couriers owned by the company; the remainder is occasional couriers crowdsourced by the company. These mixed logistics systems act as platforms that serve both retailers' delivery tasks and external tasks requested by local small businesses such as online shopkeepers, flower shops, gift shops, and take-out restaurants.

This paper studies how to schedule *internal* and *external* delivery tasks on a mixed logistics platform with multiple depots and two types of fleet, i.e., *dedicated* and *occasional* fleets. To fulfill the internal delivery tasks requested by the retailers, the dedicated vehicles owned by the logistics platform depart from the platform's depots and return to the depots after having fulfilled the deliveries. The occasional vehicles depart from their original location and pick up the goods from the depots or from external pickup locations, fulfill the delivery tasks, and finish their route at the last delivery location. Each internal delivery task must be fulfilled by either a dedicated vehicle or an occasional vehicle. However, the external delivery tasks requested by local small businesses may be undertaken or rejected by the platform due to capacity limitations or an unfavorable cost-benefit analysis. To fulfill an external delivery task requested by a shop, a dedicated or occasional vehicle travels to the shop's location to pick up the goods, and then travels to the site of the shop's customer to perform the delivery. For a logistics platform, all of the internal tasks must be fulfilled and undertaking external delivery tasks could bring extra benefit for the platform. The objective is to minimize the total traveling cost minus the extra benefit of undertaking external tasks.

The above scheduling problem represents a new variant of the close–open mixed multi-depot vehicle routing problem (VRP) with two types of delivery task and two types of vehicle fleet. The problem has the following four features: (1) multiple heterogeneous depots form the network of warehouses for the platform; (2) two types of vehicle (dedicated and occasional) constitute a mixed fleet; (3) dedicated vehicles' routes follow the structure of classical VRPs, whereas occasional vehicles adhere to open VRPs; (4) two types of task (internal and external) are fulfilled by the fleet.

2. Related works

With the emergence of ride-sharing, which offers the possibility of crowdsourced delivery to satisfy various demands at a large scale, new challenges arise in developing routing optimization models and algorithms to solve these new scheduling problems. The recent literature in this context is summarized in Alnaggar et al. (2019), who analyze current industry trends in crowdsourced delivery and provide a taxonomy of available systems based on their scheduling and matching mechanisms, target markets, and compensation schemes. Below we review works related to the problem addressed in this paper.

Arslan et al. (2019) study a variant of the dynamic pickup and delivery problem for real-time matching delivery tasks and occasional vehicles on a service platform with a rolling horizon. Qi et al. (2018) conduct an analytical study on crowdsourced delivery by considering occasional drivers' wage-response behavior in the ride-share market and investigate the optimal sizes of service zones for last-mile delivery. The core problem in these related studies on crowdsourced delivery, shared mobility, and e-commerce logistics is mainly related to studies of the classic VRP and its variants. We also study a new VRP variant.

The VRP was first investigated by Dantzig and Ramser (1959). The classic VRP usually assumes that vehicles depart from and return to a single depot and each customer is served exactly once by a vehicle. Some excellent literature reviews on the VRP have been conducted in recent years by Laporte (2009) and Toth and Vigo (2014). Among variants of the VRP, the VRP with time windows (VRPTW) generalizes the basic Capacitated VRP (CVRP) by imposing the requirement that each customer be visited within a specified time interval (a time window). Numerous scholars have conducted studies on the VRPTW, such as Desrochers et al. (1992), Bent et al. (2004), Pecin et al. (2017), and Bianchessi et al. (2019). Many exact and heuristic methods have been proposed for it, including an exact algorithm based on a set-partitioning integer formulation (Baldacci et al., 2011; Baldacci et. al., 2012), an exact algorithm based on dynamic programming and a state-space-time network (Mahmoudi and Zhou, 2016), an algorithm based on an alternating direction method of multipliers (Yao et al., 2019), an adaptive large neighborhood search (ALNS) based heuristic (Francois et al., 2019), and an exact algorithm based on column generation and cutting plane (Paradiso et al., 2020). Another related VRP variant considers both delivery tasks and pickup tasks in routing decisions. Some representative studies in this area consider new features such as divisible (or discrete split) deliveries and pickups (Nagy et al., 2015), time limits (Polat et al., 2015), and random demands and predefined customer order (Dimitrakos et al., 2015). The reader is referred to the book edited by Toth and Vigo (2014) for a comprehensive overview of exact and heuristic methods for VRPs.

Single-depot VRPs, such as the CVRP and VRPTW, do not fit the background of e-commerce logistics. E -commerce retailers usually deploy several dispersed warehouses (or distribution stations), which are generally located in urban areas. Our study is for this reason more closely related to studies that feature multiple depots in the VRP domain (Laporte et al., 1988), i.e., the multi-depot VRP (MDVRP). Salhi et al. (2013) develop a variable neighborhood search approach for the MDVRP with a fleet of heterogeneous vehicles, whereas Baldacci et al. (2013) design an exact method for the two-echelon MDVRP. Alinaghian and Shokouhi (2018) design a hybrid ALNS algorithm for solving a new MDVRP. For MDVRPs and two-echelon VRPs, the reader is referred to the excellent surveys of Montoya-Torres et al. (2015) and Cuda et al. (2015), respectively.

An open loop is a major characteristic of crowdsourced delivery–related VRPs (Qi et al., 2018). This study focuses on an open VRP (OVRP), where some of the vehicles (occasional vehicles) need not return to the depots. There have been a few studies that have combined the MDVRP and OVRP, i.e., a multi-depot open VRP (MOVRP). For example, Liu et al. (2014) develop a genetic algorithm (GA) for the MOVRP. The problem is also addressed by Lalla-Ruiz et al. (2016), who propose an improved mixed integer programming (MIP) model for the problem. Lahyani et al. (2019) develop a meta-heuristic based on ALNS to solve the MOVRP with up to six depots and 288 customers. Liu and Jiang (2012) address a close–open mixed VRP (COMVRP) involving a fleet of heterogeneous vehicles. Azadeh and Farrokhi-Asl (2019) combine the COMVRP and MDVRP and investigate a close–open mixed multi-depot VRP (COMMVRP), which represents the closest model proposed in the literature to the problem addressed in this paper. Azadeh and Farrokhi-Asl (2019) describe an MIP model for the COMMVRP and design a GA for its solution. Our work further extends the COMMVRP by addressing the problem in the context of the crowdsourced delivery industry. Indeed, in addition to the COMMVRP, we consider different types of delivery task (internal or external), different pickup locations (depots or external shops), and different starting positions for the vehicles.

2.1 Contributions of this paper

In this paper, we investigate a new variant of the COMMVRP that finds important applications in the context of the crowdsourced delivery industry. We describe MIP models for the problem and for its special case that arises when only internal deliveries are considered. To solve the problem efficiently, column generation (CG) techniques are used. Numerical experiments are performed to verify the efficiency of the proposed solution methods. The results obtained show that the proposed method can produce high-quality solutions in a limited amount of computing time for large-scale instances involving up to 20 depots, 200 internal customers, 40 external tasks, and 70 double-type vehicles. Managerial insights are reported based on the results of a sensitivity analysis.

The remainder of this paper is organized as follows. The next section describes the problem addressed in this paper in details. Section 4 proposes an MIP model as well as a CG based solution method for the problem without external deliveries. Extensions of the model to the general case of external deliveries as well as a CG based solution method are elaborated and validated in Section 5. A computational study for a large-scale instance taken from a real-world application is presented in Section 6. Finally, we conclude the paper and indicate future research directions in Section 7.

3. Problem description

We consider an e-commerce retailer running a logistics service platform as a corporate spin-off. The platform mainly fulfills delivery tasks for its parent company, but it can also serve outside businesses within its delivery capacity. Fulfilling the delivery tasks for the parent company, named *internal deliveries*, has a higher priority than fulfilling the deliveries for outside businesses, which are named *external deliveries*.

The logistics service platform has a network of warehouses or *depots*, each of which has a set of *dedicated* vehicles. The route of each dedicated vehicle is called a *dedicated route*, on which the dedicated vehicle departs from its depot and returns to it after having completed its deliveries; dedicated vehicles perform no other kinds of routes. Further, the depots are not identical with respect to the inventory of goods, and an internal delivery can potentially be supplied from more than one depot, but a customer's demand cannot be split among the different depots. Each external delivery defines both pickup and delivery locations and a vehicle servicing it must first visit the pickup location to collect the goods and then the associated delivery location.

In addition to the dedicated vehicles, a fleet of *occasional* vehicles is also available. Each occasional vehicle has a specified starting location from which its route must originate. The vehicle on an *occasional route* (i.e., a route performed by an occasional vehicle) departs from its starting location, picks up the goods from the depots or external pickup locations, fulfills the delivery tasks, and stops at the last delivery location. An internal or external delivery can be served by either a dedicated or occasional vehicle, and a dedicated or occasional vehicle can perform at most one vehicle route.

Each vehicle (dedicated or occasional) is associated with a given capacity, and in a vehicle route the total load of the deliveries served by the route must not exceed the vehicle capacity, as for the classical CVRP. Each internal delivery task must be served by exactly one vehicle, whereas an external delivery task can be left unserved, i.e., the logistics platform can undertake or reject an external delivery task according to its capacity and the benefit of delivering it. If an external delivery is served, then it must be assigned to exactly one vehicle.

The problem objective is to minimize the total cost of the routes performed by the vehicles minus the extra benefit of undertaking external tasks.

In the following, we first consider a special case of the problem considering internal deliveries only. Figure 1 shows the scheme of a solution involving internal deliveries only. In particular, the figure shows different types of vehicle route associated with the fleet of dedicated and occasional vehicles.

For this special case, in the next section we present both a mathematical formulation and a solution method based on a CG procedure. In Section 5 we then extend the mathematical formulation and the solution method to the more general case, involving both internal and external deliveries. Figure 2 shows the scheme of a solution of the general problem involving both internal and external deliveries.

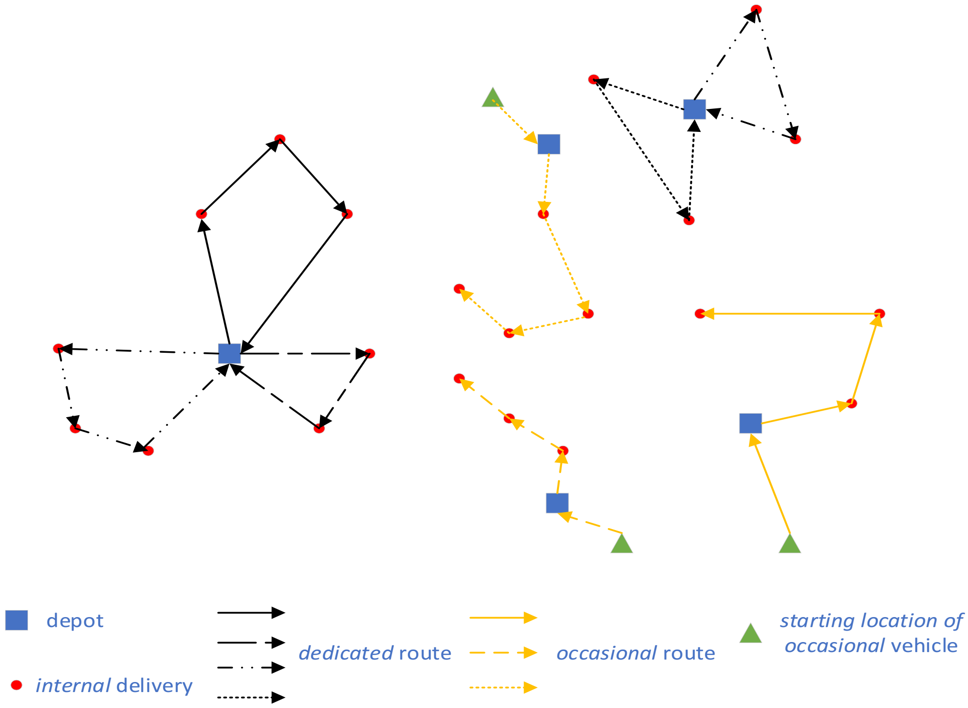


Figure 1: Example of a solution with internal deliveries only

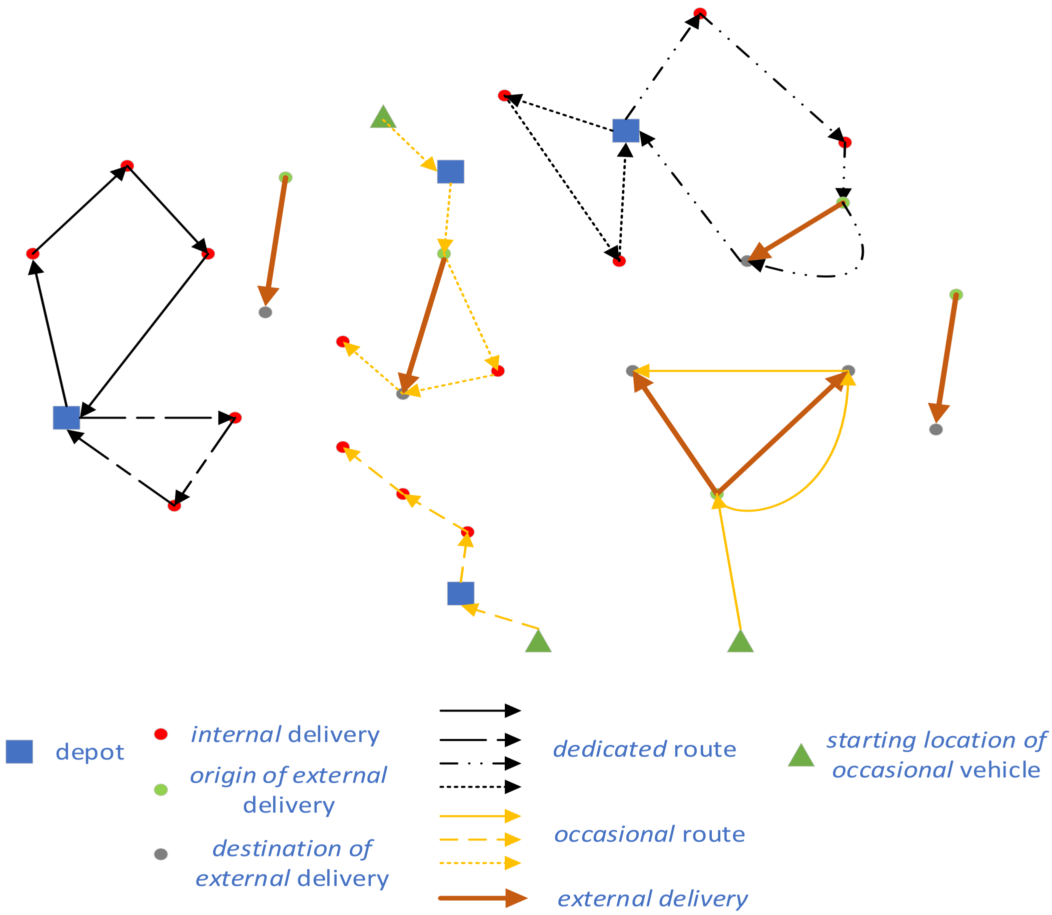


Figure 2: Example of a solution for the general problem

4. Scheduling internal deliveries

In this section, an MIP model and a CG-based solution method are designed for the special case of the problem where dedicated and occasional vehicles are used to serve internal delivery tasks only. We also present numerical experiments aimed at demonstrating the efficiency of the proposed solution method and the effectiveness of the proposed model. Managerial implications are also obtained based on a sensitivity analysis.

4.1 Mathematical formulation

In this section, an MIP model is formulated for the problem. The following notation is used in the model.

Indices and sets

K set of the dedicated vehicles, indexed by k.

K' set of the occasional vehicles, indexed by k.

D set of the depots, indexed by d.

N set of the customers (and their locations) representing the internal deliveries, indexed by i and j.

D_i set of the depots that can meet customer i's demand, indexed by d, $d \in D_i$.

N_d set of the customers whose demands can be met by depot d.

O set of the departure location of the occasional vehicles.

A set of all nodes, $A = D \cup N \cup O$.

Parameters:

q_i demand of customer i, i.e., demand associated with the internal delivery i.

t_{ij} travel time between customer i and customer j.

c_{ij} cost for dedicated vehicle k to travel from customer i to customer j.

c'_{ij} cost for occasional vehicle k to travel from customer i to customer j.

e_k load capacity of dedicated vehicle k.

e'_k load capacity of occasional vehicle k.

o_k origin of occasional vehicle k.

d_k index of the depot to which dedicated vehicle k belongs.

h_i latest service or delivery time of customer i.

w unit penalty cost of delay from the latest delivery time.

M a sufficiently large positive number.

The mathematical formulation uses the following decision variables.

α_i binary variable, which equals one if customer i is served by an occasional vehicle, and zero if customer i is served by a dedicated vehicle.

β_{id}	binary variable, which equals one if customer i's required goods are delivered from depot d, and zero otherwise. The goods are delivered by either a dedicated vehicle that belongs to the depot d, or one of all the available occasional vehicles.
γ_{ik}	binary variable, which equals one if customer i is served by dedicated vehicle k, and zero otherwise.
γ'_{ik}	binary variable, which equals one if customer i is served by occasional vehicle k, and zero otherwise.
φ_{dk}	binary variable, which equals one if occasional vehicle k travels from its origin to depot d for picking up a parcel, and zero otherwise.
δ_{ijk}	binary variable, which equals one if dedicated vehicle k visits node j immediately after visiting node i, and zero otherwise. $i, j \in N \cup D, k \in K$.
δ'_{ijk}	binary variable, which equals one if occasional vehicle k visits node j immediately after visiting node i, and zero otherwise. $i, j \in A, k \in K'$.
μ_k	nonnegative variable, departure time of dedicated vehicle k from its depot.
η_k	nonnegative variable, return time of dedicated vehicle k to its depot.
λ_{ik}	nonnegative variable, time when dedicated vehicle k visits customer i.
λ'_{ik}	nonnegative variable, time when occasional vehicle k visits customer i.

The mathematical formulation is as follows.

$$\textit{Minimize } \sum_{i\in D\cup N}\sum_{j\in D\cup N}\sum_{k\in K} c_{ij}\delta_{ijk} + \sum_{i\in A}\sum_{j\in D\cup N}\sum_{k\in K'} c'_{ij}\,\delta'_{ijk} + w\sum_{i\in N}\left(\sum_{k\in K}\lambda_{ik} + \sum_{k\in K'}\lambda'_{ik} - h_i\right)^+ \quad (1)$$

subject to:

$$\sum_{d\in D_i}\beta_{id} = 1 \qquad \forall i \in N \quad (2)$$

$$\sum_{k\in K'}\gamma'_{ik} = \alpha_i \qquad \forall i \in N \quad (3)$$

$$\delta'_{o_k dk} + 1 \geq \gamma'_{ik} + \beta_{id} \qquad \forall i \in N; k \in K'; d \in D_i; o_k \in O \quad (4)$$

$$2 \geq \sum_{j\in N_d}\delta'_{djk} + 1 \geq \gamma'_{ik} + \beta_{id} \qquad \forall i \in N; k \in K'; d \in D_i \quad (5)$$

$$2 \geq \sum_{j\in d\cup N_d}\delta'_{jik} + 1 \geq \gamma'_{ik} + \beta_{id} \qquad \forall i \in N; k \in K'; d \in D_i \quad (6)$$

$$\sum_{j\in N_d/i}\delta'_{ijk} \leq \sum_{j\in D\cup N_d/i}\delta'_{jik} = \gamma'_{ik} \qquad \forall i \in N; k \in K' \quad (7)$$

$$\delta'_{o_k dk} = \sum_{j\in N_d}\delta'_{djk} = \varphi_{dk} \qquad \forall\, k \in K'; d \in D; o_k \in O \quad (8)$$

$$\sum_{d\in D}\varphi_{dk} \leq 1 \qquad \forall\, k \in K' \quad (9)$$

$$\sum_{i\in N}\gamma'_{ik}\cdot q_i \leq e'_k \qquad \forall\, k \in K' \quad (10)$$

$$\lambda'_{ik} + t_{ij} \leq \lambda'_{jk} + M(1 - \delta'_{ijk}) \qquad \forall i \in A; j \in D \cup N; k \in K' \quad (11)$$

$$\sum_{k\in K}\gamma_{ik} = 1 - \alpha_i \qquad \forall i \in N \quad (12)$$

$$\sum_{j\in\{d_k\}\cup N_{d_k}}\delta_{jik} = \sum_{j\in\{d_k\}\cup N_{d_k}}\delta_{ijk} = \gamma_{ik} \qquad \forall i \in N; k \in K; d_k \in D_i \quad (13)$$

$$\sum_{j\in N_{d_k}}\delta_{jd_k k} = \sum_{j\in N_{d_k}}\delta_{d_k jk} \leq 1 \qquad \forall k \in K; d_k \in D \quad (14)$$

$$\sum_{i\in N}\gamma_{ik}\cdot q_i \leq e_k \qquad \forall k \in K \quad (15)$$

$$\mu_k + t_{d_k i} \le \lambda_{ik} + M(1 - \delta_{d_k ik}) \qquad \forall i \in N; k \in K; d_k \in D \qquad (16)$$

$$\lambda_{ik} + t_{ij} \le \lambda_{jk} + M(1 - \delta_{ijk}) \qquad \forall\, i,j \in N; k \in K \qquad (17)$$

$$\lambda_{ik} + t_{id_k} \le \eta_k + \ M(1 - \delta_{id_k k}) \qquad \forall i \in N; k \in K; d_k \in D \qquad (18)$$

$$\alpha_i \in \{0,1\} \qquad \forall i \in N \qquad (19)$$

$$\beta_{id} \in \{0,1\} \qquad \forall i \in N; d \in D \qquad (20)$$

$$\gamma_{ik} \in \{0,1\} \qquad \forall i \in N; k \in K \qquad (21)$$

$$\gamma'_{ik} \in \{0,1\} \qquad \forall i \in N; k \in K' \qquad (22)$$

$$\varphi_{dk} \in \{0,1\} \qquad \forall k \in K'; d \in D \qquad (23)$$

$$\delta_{ijk} \in \{0,1\} \qquad \forall i,j \in N \cup D; k \in K \qquad (24)$$

$$\delta'_{ijk} \in \{0,1\} \qquad \forall i \in A; j \in N \cup D; k \in K' \qquad (25)$$

$$\mu_k, \eta_k \ge 0 \qquad \forall i \in N;\ k \in K \qquad (26)$$

$$\lambda_{ik} \ge 0 \qquad \forall i \in N;\ k \in K \qquad (27)$$

$$\lambda'_{ik} \ge 0 \qquad \forall i \in A; k \in K'. \qquad (28)$$

Objective (1) minimizes the sum of the total travelling cost of dedicated and occasional vehicles and of the penalty costs associated with late deliveries. Constraints (2) state each customer's demand must be satisfied by one depot, from which either a dedicated vehicle or an occasional vehicle deliver the goods to the customer. The following two sets of constraints are present in the model.

Constraints on occasional vehicles. Constraints (3) connect two variables α_i and γ'_{ik} that are both related to the occasional vehicles. Constraints (4) state that if customer i is served by occasional vehicle k (i.e., $\gamma_{ik} = 1$) and customer i's goods are delivered from depot d ($\beta_{id} = 1$), vehicle k needs to travel from its origin o_k to depot d (i.e., $\delta'_{o_k dk} = 1$). Constraints (5) state that if customer i is served by the occasional vehicle k (i.e., $\gamma'_{ik} = 1$) and customer i's goods are delivered from depot d ($\beta_{id} = 1$), vehicle k travels from depot d to a customer j whose required goods are also stored in depot d (i.e., $\delta'_{djk} = 1$). Constraints (6) state that if customer i is served by occasional vehicle k and depot d, vehicle k travels from either the depot d or a customer j whose required goods are also stored in depot d to the customer i. Constraints (7) state that if customer i is served by the occasional vehicle k, there must be one predecessor, and at most one successor in its vehicle route. Constraints (8) connect two variables δ'_{ijk} and φ_{dk} that are all related to the occasional vehicles. Constraints (9) state that occasional vehicle k visits at most one depot for picking up a parcel. Constraints (10) ensure that the total loads of an occasional vehicle cannot exceed its vehicle capacity. Constraints (11) mean time flow constraint of occasional vehicle. These constraints impose both the capacity and the connectivity requirements of the solution and avoid subtours involving customers' nodes only.

Constraints on dedicated vehicles. Constraints (12) connect two variables α_i and γ_{ik} that are all related to the dedicated vehicles. Constraints (13) state that if customer i is served by a dedicated vehicle, the customer has only one predecessor and one successor, which can be either depot d that the dedicated vehicle belongs to, or a customer. Constraints (14) state that a dedicated vehicle must depart from its depot and return to the depot (if it is used). Constraints (15) ensure that the total loads of a dedicated vehicle cannot exceed its vehicle capacity. Constraints (16)~(18) are travel time conservation constraints of dedicated vehicles. Finally, Constraints (19)~(28) define the domains of the decision variables.

4.2 A CG-based solution method

Commercial solvers such as IBM-ILOG CPLEX have difficulty solving the above MIP model for large-scale problem instances, while CG techniques have been proven to be very effective in solving many large-scale problems of a difficulty no standard commercial MIP solver could cope with (Lübbecke and Desrosiers, 2005). In this section, we describe a CG-based algorithm to solve the problem that also constitutes a contribution to the literature on algorithms for solving COMMVRP variants.

4.2.1 A set-partitioning based mathematical formulation

In this section, we reformulate the problem as a master problem (MP) model by using Dantzig-Wolfe decomposition (Dantzig and Wolfe, 1960). We define $\mathcal{P}_k$ as the set of all feasible routes for vehicle k, $k \in K \cup K'$. A binary variable x_{p_k} is defined for each route $p_k \in \mathcal{P}_k$, $k \in K \cup K'$. If route p_k is chosen in solution for vehicle k, x_{p_k} equals one, zero otherwise. A binary parameter y_{i,p_k} is also used to denote whether customer i is served by route p_k. The cost for vehicle k associated with route p_k is denoted by C_{p_k}, which also includes the penalty costs of late deliveries. Based on the above definition, the set-partitioning model is formulated as follows.

[**MP**] *Minimize* $\sum_{k \in K \cup K'} \sum_{p_k \in \mathcal{P}_k} C_{p_k} x_{p_k}$ (29)

subject to:

$$\sum_{k \in K \cup K'} \sum_{p_k \in \mathcal{P}_k} y_{i,p_k} x_{p_k} = 1 \quad \forall i \in N \quad (30)$$

$$\sum_{p_k \in \mathcal{P}_k} x_{p_k} \leq 1 \quad \forall k \in K \cup K' \quad (31)$$

$$x_{p_k} \in \{0,1\} \quad \forall k \in K \cup K', \forall p_k \in \mathcal{P}_k. \quad (32)$$

Objective (29) minimizes the total costs of the routes selected in the solution. Constraints (30) guarantee that each customer is served once by a vehicle route. Constraints (31) state that each vehicle is assigned to at most one route. Constraints (32) define the domain of the decision variables.

The above formulation cannot be solved directly due the huge number of its variables corresponding to the set of all feasible routes. Hence, in practice a CG procedure solves the linear programming (LP) relaxation of the formulation by iteratively solving a restricted master problem (LR-RMP) defined over a subset $\mathcal{P}'_k \subseteq \mathcal{P}_k$ of the whole set of routes for vehicle *k*. Problem LR-RMP is defined as follows.

[**LR-RMP**] *Minimize* $\sum_{K \cup K'} \sum_{p_k \in \mathcal{P}'_k} C_{p_k} x_{p_k}$ (33)

subject to:

$$\sum_{k \in K \cup K'} \sum_{p_k \in \mathcal{P}'_k} y_{i,p_k} x_{p_k} = 1 \quad \forall i \in N \quad (34)$$

$$\sum_{p_k \in \mathcal{P}'_k} x_{p_k} \leq 1 \quad \forall k \in K \cup K' \quad (35)$$

$$x_{p_k} \geq 0 \quad \forall k \in K \cup K', \forall p_k \in \mathcal{P}'_k. \quad (36)$$

Following the scheme of a CG procedure, at each iteration of the algorithm the optimal dual variables associated with LR-RMP are given as input to the pricing problem (PP) to generate new routes. The CG procedure terminates whenever no negative reduced cost columns or routes are generated, and the resulting LR-RMP solution corresponds to the optimal solution of the LP-relaxation of the set-partitioning based formulation. Otherwise, the columns identified by solving problem PP are added to problem LR-RMP, and a new iteration is executed. The initial sets of routes $\{\mathcal{P}_k'\}$ must be defined to guarantee the existence of an initial feasible LR-RMP solution. In our implementation, the initial set of routes is generated by using a greedy heuristic.

4.2.2 Pricing problem PP

Problem PP aims to generate columns or routes having negative reduced costs, which are then added to problem LR-MRP in an iterative fashion. At each iteration of the CG procedure, there are $|K|$+$|K'|$ number of problems to be solved. A problem PP is associated with a vehicle k and generates a feasible routing plan for the vehicle. The PP models for each vehicle $k \in K \cup K'$ (denoted as PP_k and $PP_{k'}$) are defined as follows. Let π_i and ω_k be the dual variables associated with constraints (34) and (35), respectively.

The mathematical formulations of problem PP use the following decision variables:

$\theta_{i,j}$ binary variable, which equals one if vehicle $k \in K \cup K'$ visits node j immediately after node i, and zero otherwise.

τ_i binary variable, which equals one if vehicle $k \in K \cup K'$ visits customer i, and zero otherwise.

ζ nonnegative variable, departure time of dedicated vehicle $k \in K$ from its depot.

ψ nonnegative variable, return time of dedicated vehicle $k \in K$ to its depot.

ρ_d binary variable, which equals one if occasional vehicle $k' \in K'$ visits the depot d, and zero otherwise.

ξ_i nonnegative variable, time when dedicated vehicle $k \in K$ visits customer i.

ξ_i' nonnegative variable, time when occasional vehicle $k' \in K'$ visits customer i.

The mathematical formulations PP_k and $PP_{k'}$ are as follows.

(1) The PP for an occasional vehicle $k' \in K'$

[PP$_{k'}$] *Minimize* $\sigma_{k'} = C_{p_{k'}} - \sum_{i \in N} \tau_i \pi_i - \omega_{k'}$ (37)

subject to:

$\sum_{j \in N/\{i\}} \theta_{ij} \leq \sum_{j \in D \cup N/i} \theta_{ji} = \tau_i \qquad \forall i \in N$ (38)

$\theta_{o_{k'}d} = \sum_{i \in N_d} \theta_{di} = \rho_d \qquad \forall d \in D$ (39)

$\sum_{d \in D} \rho_d = 1$ (40)

$\sum_{i \in N} q_i \tau_i \leq e_{k'}$ (41)

$\xi_i' + t_{ij} \leq \xi_j' + M(1 - \theta_{ij}) \qquad \forall i \in A; j \in D \cup N$ (42)

$C_{p_{k'}} = \sum_{i \in A} \sum_{j \in D \cup N} c_{ij}' \theta_{ij} + w \sum_{i \in N} (\xi_i' - h_i)^+$ (43)

$\tau_i \leq \rho_d \qquad \forall d \in D; i \in N_d$ (44)

$\tau_i \in \{0,1\} \qquad \forall i \in N$ (45)

$\theta_{ij} \in \{0,1\} \qquad \forall i \in A; j \in N$ (46)

$\rho_d \in \{0,1\} \qquad \forall d \in D$ (47)

$$\xi_i' \geq 0 \qquad \forall i \in A. \tag{48}$$

Objective (37) minimizes the reduced cost. Constraints (38)~(42) correspond to Constraints (7)~(11), respectively. Constraints (43) represent the calculation of the cost of a column, which is the sum of the occasional vehicle's delivery costs and of the penalty costs. Constraints (44) ensure that an occasional vehicle should go to a depot d that contains the goods required by customer i. Constraints (45)~(48) define the decision variables.

(2) The PP for a dedicated vehicle $\boldsymbol{k} \in \boldsymbol{K}$

[**PP$_k$**] *Minimize* $\sigma_k = C_{p_k} - \sum_{i \in N} \tau_i \pi_i - \omega_k$ (49)

subject to:

$$\sum_{j \in N_{d_k}} \theta_{d_k j} = 1 \tag{50}$$

$$\sum_{i \in N_{d_k}} \theta_{i d_k} = 1 \tag{51}$$

$$\sum_{j \in \{d_k\} \cup N_{d_k}} \theta_{ij} = \sum_{j \in N_{d_k} \cup \{d_k\}} \theta_{ji} = \tau_i \qquad \forall i \in N \tag{52}$$

$$\sum_{i \in N} q_i \tau_i \leq e_k \tag{53}$$

$$\zeta + t_{d_k i} \leq \xi_i + M(1 - \theta_{d_k i}) \qquad \forall i \in N \tag{54}$$

$$\xi_i + t_{ij} \leq \xi_j + M(1 - \theta_{ij}) \qquad \forall i \in N; j \in N \tag{55}$$

$$\xi_i + t_{i d_k} \leq \psi + M(1 - \theta_{i d_k}) \qquad \forall i \in N \tag{56}$$

$$C_{p_k} = \sum_{i \in \{d_k\} \cup N} \sum_{j \in N \cup \{d_k\}} c_{ij} \theta_{ij} + w \sum_{i \in N} (\xi_i - h_i)^+ \tag{57}$$

$$\tau_i \in \{0,1\} \qquad \forall i \in N \tag{58}$$

$$\theta_{ij} \in \{0,1\} \qquad \forall i, j \in N \tag{59}$$

$$\zeta, \xi_i, \psi \geq 0 \qquad \forall i \in N. \tag{60}$$

Objective (49) minimizes the reduced cost. Constraints (50)~(56) correspond to Constraints (12)~(18), respectively. Constraint (57) represents the calculation of the cost of a column. Constraints (58)~(60) define the domain of the decision variables.

4.2.4 A column generation-based heuristic

The optimal solution cost of formulation LR-RMP, computed as described in the previous section using the column generation method, provides a valid lower bound on the optimal solution cost of the problem. If the LR-RMP solution is also an integer solution, then the solution represents the optimal solution of the problem. Generally speaking, for difficult instances of the problem, the optimal solution of LR-RMP is generally fractional, i.e., some of the variables x_{p_k} are selected in solution with fractional values in the interval (0,1). To produce an integer (and feasible) solution to the problem, we design a variable-fixing algorithm based on the procedure proposed by Wang et al. (2018), which consists in heuristically fixing the value of selected x_{p_k} variables to 1. The heuristic works as follows.

We use two *resources,* $customer_i$ for all $i \in N$ (*customer resource*) and $vehicle_k$ for all $k \in K \cup K'$ (*vehicle resource*), that are associated with the right-hand sides of constraints (34) and (35) in problem LR-RMP, respectively. A resource equal to 0 means that the corresponding constraint is no longer active or, in other words, that a customer has been already served or that a vehicle has been already used in the solution. Hence, after fixing the resources we obtain a new LR-RMP problem that can be used to generate new columns considering the new dual values in the pricing problem. At each iteration of the algorithm, set Ω is used to represent the set of routes selected in the partial solution, while at the end of the algorithm the set contains the routes selected in solution. In this context, the solution approach consists of the following steps.

Step 0. Initialize the resources $customer_i = 1$ for all $i \in N$, and $vehicle_k = 1$ for all $k \in K \cup K'$. Initialize the set Ω as the empty set.

Step 1. Solve problem LR-RMP in a column generation fashion, where the right-hand sides of constraint (34) and (35) are defined according to the values of the resources. Let $\bar{\mathcal{P}}$ be the set of routes selected in the optimal LR-RMP solution, i.e., the set of routes such that the corresponding variables x_{p_k} are nonnegative.

Step 2. Select from set $\bar{\mathcal{P}}$ all those variables or routes that are equal to 1, plus the fractional variable of highest value; in case of ties, the variable having the lower cost C_{p_k} is selected. Update set Ω with the selected routes.

Step 3. Update the resources based on the set of routes $\bar{\mathcal{P}}$, i.e., all resources corresponding to customers served and to vehicles used in the routes of set $\bar{\mathcal{P}}$ are set equal to 0.

Step 4. If $customer_i = 0$ for all $i \in N$, then terminate the algorithm and an integer solution represented by the routes in set Ω has been computed. Otherwise, go to Step 1.

The above procedure can always find a feasible solution if each customer can be served by at least a single vehicle route and the number of vehicles available is sufficiently large.

The flowchart of the CG-based solution method is shown in Figure 3.

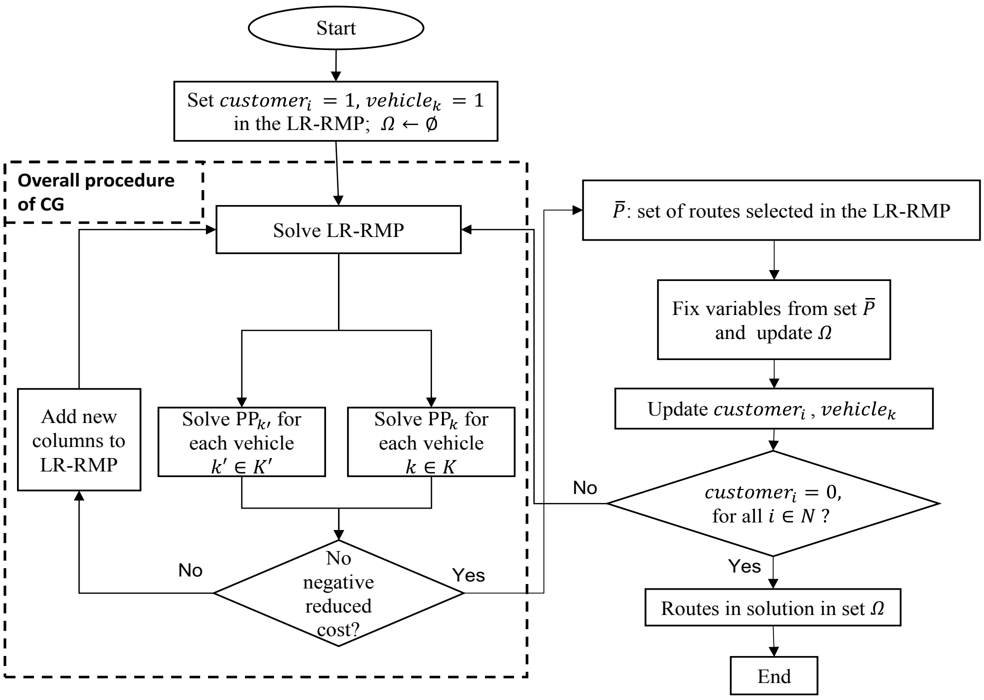


Figure 3: Flowchart of the CG-based solution method

4.3 Computational study

In this section, we present numerical experiments to validate the efficiency of the proposed solution method and the effectiveness of the proposed model. Experiments were conducted on a computer equipped with an Intel(R) Core (TM) i7-8565U CPU, 1.8 GHz processing speed, and 20 GB of memory. The algorithm was implemented in C# (Visual Studio 2019), and the general-purpose MIP solver IBM-ILOG CPLEX version 12.6.1 was used to solve the LR-RMP and the PP models.

4.3.1 Test instances

We use seven randomly generated instance groups (ISGs) to investigate the performance of the algorithm and to conduct a sensitivity analysis. The dimensions of the seven ISGs are shown in Table 1. In addition, the dedicated vehicles' capacity is set equal to 7 whereas the occasional vehicles' capacity ranges in the interval [5,8]. The demand of customers ranges in the interval [1,2]. The unit penalty cost of delay from the latest delivery time is set equal to 1. The latest service or delivery time of customers ranges in the interval [0.5,6.5].

Table 1: Data of the test instances

	depot $\|D\|$	dedicated vehicle $\|K\|$	occasional vehicle $\|K'\|$	total number of vehicles $\|K\|+\|K'\|$	customer points $\|N\|$
ISG1	2	4	2	6	12
ISG2	3	4	3	7	15
ISG3	3	6	2	8	18
ISG4	3	6	4	10	24
ISG5	4	8	3	11	30
ISG6	4	8	5	13	42
ISG7	4	8	6	14	60

4.3.2 Evaluating the efficiency of the CG based solution method

To evaluate the performance of the CG-based solution method, we compare the results obtained by the CG method with the results obtained by CPLEX on the small-scale instance groups ISG1, ISG2, and ISG3. The results are given in Table 2. The table shows that CPLEX failed to solve some instances of group ISG3 within two hours of computing time. Further, the table shows that for the instances that were solved by CPLEX, CG performs quite well, with an average gap value equal to 0.53%, as shown by the column "Gap" under the CG section of Table 2.

4.3.3 Evaluating the effectiveness of the proposed model

To further evaluate the effectiveness of the proposed model, we compare the results obtained by solving our model and the results obtained by some intuitive but practical decision rules. The comparative experiments are conducted for the medium-scale groups ISG4, ISG5, ISG6, and ISG7. The details of the decision rule used in these comparative experiments are as follows.

Decision makers in a real-world environment may not be familiar with the methodology of mathematical programming and make their decisions based on some quality rules. They may select the vehicles one by one according to some basic rule or order. In this study, two types of order are considered: the "capacity-cost" order and the "cost-capacity" order.

Table 2: Comparison between CPLEX and CG on small-sized instances

Instances		CPLEX		CG		Gap
Group	number	F_{Cplex}	t_{Cplex}(s)	F_{CG}	t_{CG}(s)	
ISG1	1	15.65	7	15.65	11	0.00%
	2	19.58	2	19.58	15	0.00%
	3	15.82	8	15.98	12	1.01%
	4	13.06	7	13.31	13	1.91%
	5	13.63	6	13.81	9	1.32%
	6	17.28	12	17.28	13	0.00%
	7	20.20	10	20.28	15	0.40%
	8	15.75	6	15.88	14	0.83%
ISG2	9	21.10	674	21.23	23	0.62%
	10	18.51	90	18.53	20	0.11%
	11	20.78	634	20.78	21	0.00%
	12	17.36	409	17.36	33	0.00%
	13	18.46	615	18.46	31	0.00%
	14	17.20	284	17.28	22	0.47%
	15	19.03	272	19.03	21	0.00%
	16	19.44	87	19.44	28	0.00%
ISG3	17	19.97	5663	23.83	40	2.30%
	18	——	——	24.69	41	——
	19	——	——	22.79	34	——
	20	——	——	23.83	40	——
Average						0.53%

Notes: F_{Cplex}, F_{CG} represent the solution costs computed by CPLEX and CG, respectively; t_{Cplex}, t_{CG} represent the computation time of CPLEX and CG, respectively. "Gap" is the percentage gap between the optimal solution obtained by CPLEX and the solution obtained by the CG method.

In the "capacity-cost" order, vehicles are sorted according to decreasing values of vehicle capacity, and ties are broken by selecting the vehicle with the lower transportation cost. The "cost-capacity" order adopts a similar rule, but the transportation costs are used as the first criterion, and vehicle capacity as the second. Customers are assigned to vehicles using a greedy heuristic following a nearest-insertion rule. For a problem instance, we select the rule resulting in the lower value of the objective function. Table 3 reports the results of a comparison between the two rules and the CG method. It clearly shows that the CG method can achieve significant cost savings with respect to quality rules (on average about 52%). This significant gap demonstrates the advantage of the operations research (OR) methodology in the context of this problem.

For the medium-scale instances, Table 3 shows that the CG method can solve these instances within two hours of computing time, thus being very effective in comparison with the results of commercial solvers such as CPLEX.

Table 3: Comparison between the rule-based approach and CG on medium-sized instances

Instances		Model		Rule	Gap
Group	number	F_{CG}	t_{CG}(s)	F_{DR}	
ISG4	1	25.65	119	37.96	47.99%
	2	25.53	150	36.66	43.60%
	3	25.86	156	46.22	78.73%
	4	27.52	147	37.42	35.97%
ISG5	5	38.22	261	51.99	36.03%
	6	31.09	212	56.42	81.47%
	7	36.43	191	56.50	55.09%
	8	29.95	272	48.10	60.60%
ISG6	9	41.87	786	66.39	58.56%
	10	39.98	848	59.56	48.97%
	11	37.05	1053	61.50	65.99%
	12	44.01	843	68.92	56.60%
ISG7	13	59.50	5247	81.77	37.43%
	14	58.44	5681	89.10	52.46%
	15	55.04	4667	86.29	56.78%
	16	65.71	4898	79.47	20.94%
Average					52.33%

Notes: F_{DR}, F_{CG} represent the objective values obtained by the decision rules (DR) and method CG, respectively; t_{CG} gives the computation time of CG. "Gap" reports percentage gap between the decision rules and CG.

5. Scheduling internal and external deliveries

As mentioned in the introduction, an e-commerce logistics service platform may not only serve the deliveries of its parent company (the e-commerce retailer), but also perform deliveries for local small businesses that cannot maintain their own logistics network (external deliveries), such as flower shops, gift shops, and take-out restaurants. In this section, we extend the models and the solution approach described for the special case of internal delivery only to the more general case where both internal and external deliveries are considered. We also present numerical experiments to validate the extended model and solution method.

5.1 A mathematical formulation for the general case

In this section, we revise the mathematical formulation introduced in Section 4 for the internal deliveries case to include the external deliveries. We rewrite the objective function of the problem to add a term that considers the benefits of fulfilling external deliveries as part of a solution.

The following notation is used in addition to the notation already introduced in Section 4:

L set of the external delivery tasks, indexed by l.

$g_l, \bar{g}_l$ index of origin and destination of the external delivery task l.

G' set of all origins of the external delivery tasks; $G' = \{g_l | l \in L\}$.

G'' set of all destinations of the external delivery tasks; $G'' = \{\bar{g}_l | l \in L\}$.

G set of all origins and destinations of the external delivery tasks; $G = \{g_l, \bar{g}_l | l \in L\}$.

A set of all nodes, $A = D \cup N \cup O \cup G$.

p_l demand of external delivery task l.

v_l revenue for fulfilling the external delivery task l.

The new mathematical formulation uses, in addition to the decision variables introduced in Section 4, a binary variable ϕ_l that is equal to one if external delivery task l is undertaken, zero otherwise. The new mathematical formulation is as follows.

The objective function (1) is revised as objective (61) by considering the revenue earned by undertaking external deliver tasks, i.e., $\sum_{l \in L} v_l \, \phi_l$.

$$Minimize \sum_{i \in D \cup N \cup G} \sum_{j \in D \cup N \cup G} \sum_{k \in K} c_{ij} \delta_{ijk} + \sum_{i \in A} \sum_{j \in D \cup N \cup G} \sum_{k \in K'} c'_{ij} \, \delta'_{ijk} + w \sum_{i \in N \cup G''} (\sum_{k \in K} \lambda_{ik} + \sum_{k \in K'} \lambda'_{ik} - h_i)^+ - \sum_{l \in L} v_l \phi_l \qquad (61)$$

Constraints (4), (10), and (15) are revised as constraints (62), (63), and (64), respectively.

$$2 \geq \sum_{j \in \{d\} \cup G'} \delta'_{o_k jk} + 1 \geq \gamma'_{ik} + \beta_{id} \qquad \forall i \in N; k \in K'; d \in D_i; o_k \in O \qquad (62)$$

$$\sum_{i \in N} \gamma'_{ik} \cdot q_i + \sum_{l \in L} \sum_{\bar{g}_l \in G''} \gamma'_{\bar{g}_l k} \cdot p_l \leq e'_k \qquad \forall \, k \in K' \qquad (63)$$

$$\sum_{i \in N} \gamma_{ik} \cdot q_i + \sum_{l \in L} \sum_{\bar{g}_l \in G''} \gamma_{\bar{g}_l k} \cdot p_l \leq e_k \qquad \forall k \in K. \qquad (64)$$

Further, in order to consider the external deliveries, in constraints (5)~(8), (11), (13), (14) and (25), the set to which the index j belongs is extended to include set G; also in constraints (21)~(22), and (27), the set to which the index i belongs is extended to include set G; in constraints (16), the set to which the index i belongs is extended to include set G'; in constraints (17) and (24), the set to which the index i and j belong is extended with set G; in constraints (18), the set to which the index i belongs is extended with set G''. Besides the above modification on constraints, the following new constraints are considered.

$$\sum_{j \in \{o_k\} \cup G} \delta'_{jdk} = \sum_{j \in N_d \cup G} \delta'_{djk} \qquad \forall k \in K'; d \in D \qquad (65)$$

$$\sum_{j \in D \cup N \cup G} \delta'_{g_l jk} = \sum_{j \in \{o_k\} \cup D \cup N \cup G / \{\overline{g_l}\}} \delta'_{j g_l k} = \gamma'_{g_l k} \qquad \forall l \in L; g_l \in G'; k \in K'; o_k \in O \qquad (66)$$

$$\sum_{j \in D \cup N \cup G / \{g_l\}} \delta'_{\bar{g}_l jk} \leq \sum_{j \in D \cup N \cup G} \delta'_{j \bar{g}_l k} = \gamma'_{\bar{g}_l k} \qquad \forall l \in L; \bar{g}_l \in G''; k \in K' \qquad (67)$$

$$\gamma'_{g_l k} = \gamma'_{\bar{g}_l k} \qquad \forall l \in L; g_l \in G'; \bar{g}_l \in G''; k \in K' \qquad (68)$$

$$\lambda'_{g_l k} \leq \lambda'_{\bar{g}_l k} \qquad \forall l \in L; g_l \in G'; \bar{g}_l \in G''; k \in K' \qquad (69)$$

$$\sum_{j \in \{d_k\} \cup N_{d_k} \cup G / \{\bar{g}_l\}} \delta_{j g_l k} = \sum_{j \in N_{d_k} \cup G} \delta_{g_l jk} = \gamma_{g_l k} \qquad \forall l \in L; g_l \in G'; k \in K; d_k \in D \qquad (70)$$

$$\sum_{j \in N_{d_k} \cup G} \delta_{j \bar{g}_l k} = \sum_{j \in \{d_k\} \cup N_{d_k} \cup G / \{g_l\}} \delta_{\bar{g}_l jk} = \gamma_{\bar{g}_l k} \qquad \forall l \in L; \bar{g}_l \in G''; k \in K; d_k \in D \qquad (71)$$

$$\gamma_{g_l k} = \gamma_{\bar{g}_l k} \qquad \forall l \in L; g_l \in G'; \bar{g}_l \in G''; k \in K \qquad (72)$$

$$\lambda_{g_l k} \leq \lambda_{\bar{g}_l k} \qquad \forall l \in L; g_l \in G'; \bar{g}_l \in G''; k \in K \qquad (73)$$

$$\phi_l = \sum_{k \in K'} \gamma'_{\bar{g}_l k} + \sum_{k \in K} \gamma_{\bar{g}_l k} \qquad \forall l \in L; \bar{g}_l \in G'' \qquad (74)$$

Constraints (65) state that if depot d is visited by occasional vehicle k, there must be one predecessor and one successor. Constraints (66) ensure that if the origin of the external delivery task g_l is served by occasional vehicle k, there must be one predecessor, and one successor; similar constraints are imposed for the destination of the external delivery task $\bar{g}_l$ by constraints (67). Constraints (68) state that if the destination of the external delivery task $\bar{g}_l$ is served by occasional vehicle k, the origin of this external delivery task g_l must be served by vehicle k. Constraints (69) ensure that occasional vehicle k must travel to the origin of the external delivery task g_l to pick up the goods before it visits the destination $\bar{g}_l$ of the external delivery task. Constraints (70) guarantee that if the origin of the external delivery task g_l is served by dedicated vehicle k, there must be one predecessor, and one successor; similarly, constraints (71) are set for the destination of the external delivery task $\bar{g}_l$. Constraints (72) state that if the destination of the external delivery task $\bar{g}_l$ is served by dedicated vehicle k, the origin of this external delivery task g_l must be served by vehicle k. Constraints (73) ensure that dedicated vehicle k must travel to the origin of the external delivery task g_l to pick up goods before it visits the destination of this external delivery task $\bar{g}_l$. Constraints (74) link constraints among variables γ_{ik}, γ'_{ik} and ϕ_l.

5.2 Extending the CG-based solution method to the general case

In this section, we describe the changes required to the CG-based solution to handle the general case where both internal and externa deliveries are considered. Below, we summarize the list of changes of the LR-RMP and the pricing problem.

(i) The following constraints (75) are added to the LR-RMP:

$$\sum_{k\in K\cup K'}\sum_{p_k\in\mathcal{P}'_k} y_{\bar{g}_l p_k} x_{p_k} \le 1 \qquad \forall l\in L;\bar{g}_l\in G'' \tag{75}$$

(ii) For the pricing problem PP$_{k'}$, the objective function (37) is replaced by the following objective:

$$Minimize\ \sigma_{k'} = C_{p_{k'}} - \sum_{i\in N\cup G''}\tau_i\pi_i - \omega_{k'} \tag{76}$$

Further, constraints (41) and (43) are revised as constraints (77) and (78), respectively:

$$\sum_{i\in N} q_i\tau_i + \sum_{l\in L}\sum_{\bar{g}_l\in G''} p_l\tau_{\bar{g}_l} \le e_{k'} \tag{77}$$

$$C_{p_{k'}} = \sum_{i\in A}\sum_{j\in D\cup N\cup G} c'_{ij}\theta_{ij} - \sum_{l\in L} v_l\tau_{\bar{g}_l} + w\sum_{i\in N\cup G''}(\xi'_i - h_i)^+ \tag{78}$$

For constraints (38)~(39), (42), and (46), the set to which the index j belongs is extended to include set G; constraints (45) are revised similarly. Handling external deliveries into the model also require the following new constraints:

$$\sum_{j\in D\cup N\cup G}\theta_{g_l j} \le \sum_{j\in\{o_{k'}\}\cup D\cup N\cup G/\{\bar{g}_l\}}\theta_{j g_l} = \tau_{g_l} \qquad \forall l\in L; g_l\in G' \tag{79}$$

$$\sum_{j\in D\cup N\cup G/\{g_l\}}\theta_{\bar{g}_l j} \le \sum_{j\in D\cup N\cup G}\theta_{j\bar{g}_l} = \tau_{\bar{g}_l} \qquad \forall l\in L;\bar{g}_l\in G'' \tag{80}$$

$$\tau_{g_l} = \tau_{\bar{g}_l} \qquad \forall l\in L; g_l\in G'; \bar{g}_l\in G'' \tag{81}$$

$$\xi'_{g_l} \le \xi'_{\bar{g}_l} \qquad \forall l\in L; g_l\in G'; \bar{g}_l\in G'' \tag{82}$$

$$\xi'_d \le \xi'_i \qquad \forall d\in D; \forall i\in N \tag{83}$$

$$\sum_{j\in D\cup G'}\theta_{o_{k'}j} = 1. \tag{84}$$

Constraints (79)~(82) correspond to constraints (66)~(69), respectively. Constraints (83) guarantee that occasional vehicle k must travel to depot d to pick up goods before it visits the customers. Constraints (84) ensure that the occasional vehicle must travel from its origin if the vehicle is used in solution.

(iii) For the pricing problem PP$_k$, objective (49) is revised as follows:

$$Minimize\ \sigma_k = C_{p_k} - \sum_{i\in N\cup G''}\tau_i\pi_i - \omega_k. \tag{85}$$

In addition, constraints (53) and (55) are revised as:

$\sum_{i\in N} q_i\tau_i + \sum_{l\in L}\sum_{\bar{g}_l\in G''} p_l\tau_{\bar{g}_l} \le e_k$ (86)

$C_{p_k} = \sum_{i\in\{d_k\}\cup N\cup G}\sum_{j\in\{d_k\}\cup N\cup G} c_{ij}\theta_{ij} - \sum_{l\in L} v_l\tau_{\bar{g}_l} + w\sum_{i\in N\cup G''}(\xi_i - h_i)^+$ (87)

Further, constraints (50)~(52), (55), and (58)~(60) are extended to include set G, and constraints (54) and (56) are also extended to include sets G' and G''. Besides, the following new constraints are added to the model:

$\sum_{j\in N_{d_k}\cup G}\theta_{g_l j} = \sum_{j\in\{d_k\}\cup N_{d_k}\cup G/\{\bar{g}_l\}}\theta_{j g_l} = \tau_{g_l}$ $\quad\forall l\in L; g_l\in G'$ (88)

$\sum_{j\in\{d_k\}\cup N_{d_k}\cup G/\{g_l\}}\theta_{\bar{g}_l j} = \sum_{j\in N_{d_k}\cup G}\theta_{j\bar{g}_l} = \tau_{\bar{g}_l}$ $\quad\forall l\in L; \bar{g}_l\in G''$ (89)

$\tau_{g_l} = \tau_{\bar{g}_l}$ $\quad\forall l\in L; g_l\in G'; \bar{g}_l\in G''$ (90)

$\xi_{g_l} \le \xi_{\bar{g}_l}$ $\quad\forall l\in L; g_l\in G'; \bar{g}_l\in G''.$ (91)

Constraints (88)~(91) correspond to constraints (70)~(73), respectively.

5.3 Computational study

5.3.1 Test instances

In this section, we report a computational study on a new group of four instances that also include external deliveries. Table 4 gives the details of the new group of instances. The aim of this section is to compare CPLEX and the rule-based approaches with the CG-based method on small and medium-size instances. In Section 6, we further extend this computational study by addressing large-scale instances involving up to 200 internal and 40 external deliveries.

Table 4: Data of the test instances

	depot numbers $\lvert D\rvert$	dedicated vehicle numbers $\lvert K\rvert$	occasional vehicle numbers $\lvert K'\rvert$	total vehicle numbers $\lvert K\rvert + \lvert K'\rvert$	customer points $\lvert N\rvert$	external delivery tasks $\lvert L\rvert$
ISG8	2	2	2	4	10	2
ISG9	2	4	2	6	12	2
ISG10	3	6	2	8	18	4
ISG11	3	6	4	10	24	5

The results of the comparison with CPLEX are reported in Table 5. The table shows that high-quality solutions can be produced by the CG method, as shown by the average optimality gap of CG (0.11%). Further, all the instances of ISG9 are solved by CG in a small fraction of the computing time of CPLEX.

To further attest the quality of the CG method, we compare it with a rule-based method like the one described in Section 4.3.3, where external tasks are considered together with internal tasks, but with a lower priority. Table 6 summarizes the comparison. The table shows that the average gap is equal to 158.79%, thus showing the difficulty of these problems, and the importance of addressing them with mathematical tools such as the method proposed in this work.

Table 5: Comparison between CG and with CPLEX on small-sized instances

Instances		CPLEX		CG		Gap
Scale	index	F_{Cplex}	t_{Cplex}(s)	F_{CG}	t_{CG}(s)	
ISG8	1	6.04	6	6.04	10	0.00%
	2	11.51	2	11.51	14	0.00%
	3	8.71	7	8.71	18	0.00%
	4	13.73	12	13.74	10	0.07%
	5	11.27	8	11.36	9	0.80%
ISG9	6	20.03	1498	20.03	33	0.00%
	7	9.57	224	9.59	25	0.21%
	8	9.04	340	9.04	21	0.00%
	9	7.65	374	7.65	24	0.00%
	10	12.75	490	12.75	20	0.00%
Average						0.11%

Notes: F_{Cplex}, F_{CG} represent the solution costs obtained by CPLEX, and CG, respectively. t_{Cplex}, t_{CG} are the computation times in seconds of CPLEX and CG, respectively. "Gap" is the optimality gap between the solution obtained by CPLEX and the solution obtained by the CG method.

Table 6: Comparison between CG and the rule-based approach on medium-sized instances

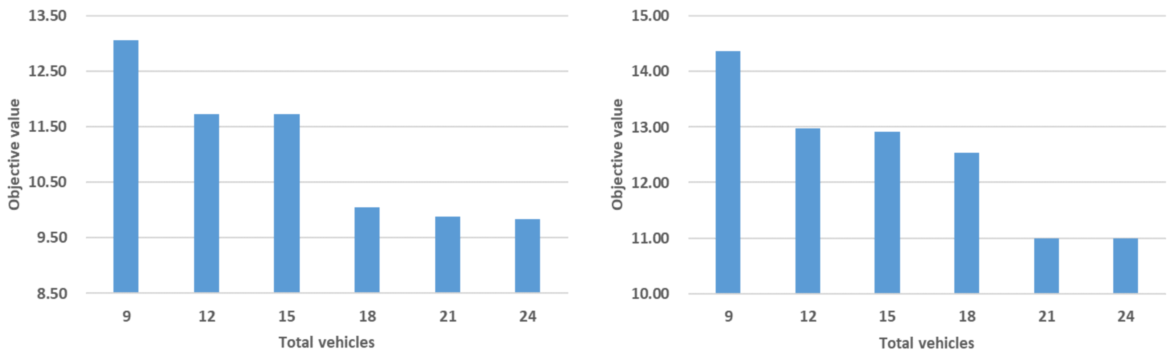
Instances		Model		Rule	Gap
scale	index	F_{CG}	t_{CG}(s)	F_{DR}	
ISG10	1	16.22	118	34.89	115.10%
	2	12.22	143	32.83	168.66%
	3	16.78	159	34.10	103.22%
	4	17.48	130	37.36	113.73%
	5	10.29	138	30.50	196.40%
ISG11	6	14.85	323	40.34	171.65%
	7	16.89	417	41.21	143.99%
	8	18.17	593	48.86	168.90%
	9	14.25	1162	40.66	185.33%
	10	17.60	572	56.48	220.91%
Average					158.79%

Notes: F_{DR}, F_{CG} represent the solution costs obtained by decision rule (DR) and by method CG, respectively; t_{CG} represents the computation time in seconds of CG. "Gap" gives the percentage gap between the decision rule and the model solved by the CG.

5.3.2 Sensitivity analysis

(1) Sensitivity analysis on the total number of vehicles

The first series of experiments is a sensitivity analysis of the total number of vehicles, which assumes a fixed ratio of 2:1 between the numbers of dedicated and occasional vehicles. The results are summarized in Figure 4 for the same instance used in the previous experiments. The x-axis denotes the total number of vehicles. From the two charts in Figure 4, we can see that the larger the total number of vehicles, the lower the total cost, no matter the value of the vehicles' unit transportation cost (i.e., c_{ij} and c'_{ij}).

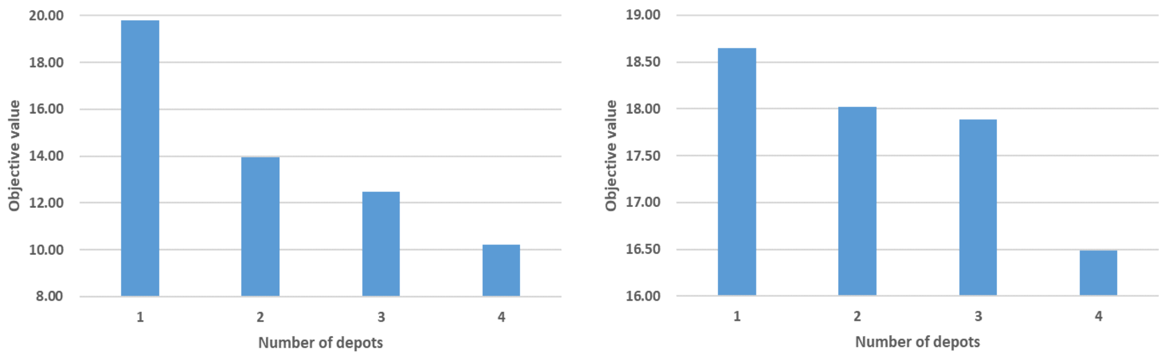


Dedicated vehicles's cost is higher than occassional vehicles ($c_{ij} > c'_{ij}$) Dedicated vehicles's cost is lower than occassional vehicles ($c_{ij} < c'_{ij}$)

Figure 4: Sensitivity analysis of the total number of vehicles

(2) Sensitivity analysis of the number of depots

The second series of experiments is a sensitivity analysis of the number of depots given the number of dedicated vehicles, occasional vehicles, and customers. For the case "$c_{ij} > c'_{ij}$", the numbers of dedicated vehicles, occasional vehicles, customers and the external delivery tasks are 6, 2, 18 and 4, respectively; while for the case "$c_{ij} < c'_{ij}$", the numbers are 6, 4, 24 and 5. The results obtained are shown in Figure 5. The results show that the larger the number of depots is, the lower the total cost is no matter which type of vehicles' unit transportation cost (i.e., c_{ij} and c'_{ij}) is adopted. If more depots are available, then there is a higher chance to assign the dedicated vehicles to reduce the total cost for fulfilling the customers' orders.



Dedicated vehicles's cost is higher than occassional vehicles ($c_{ij} > c'_{ij}$) Dedicated vehicles's cost is lower than occassional vehicles ($c_{ij} < c'_{ij}$)

Figure 5: Sensitivity analysis of the number of depots

(3) Sensitivity analysis of the regions of customers and the depots

The third series of experiments is a sensitivity analysis of the size of regions where the customers and depots are distributed. The experiments are based on the instance group ISG 11, i.e., three depots, six dedicated vehicles, four occasional vehicles, 24 customers and 5 external delivery tasks. By following the experimental setting in the study of Bertsimas et al. (2019), we assume that the 24 customers and 5 external delivery tasks are uniformly distributed in a circle with radius r_1 and the three depots are evenly located in a circle with radius r_2, both centred at the same location.

The left part of Figure 6 shows the results with different r_2 values given r_1 of five; while the right part of Figure 6 shows the results under different r_1 values given r_2 of three. The results show that given a value of a radius, there is a value for the other type of radius that minimizes the total cost. This result attests the importance of determining these radius (or planning these regions' sizes in reality) and can be potentially useful for managers in many ways.

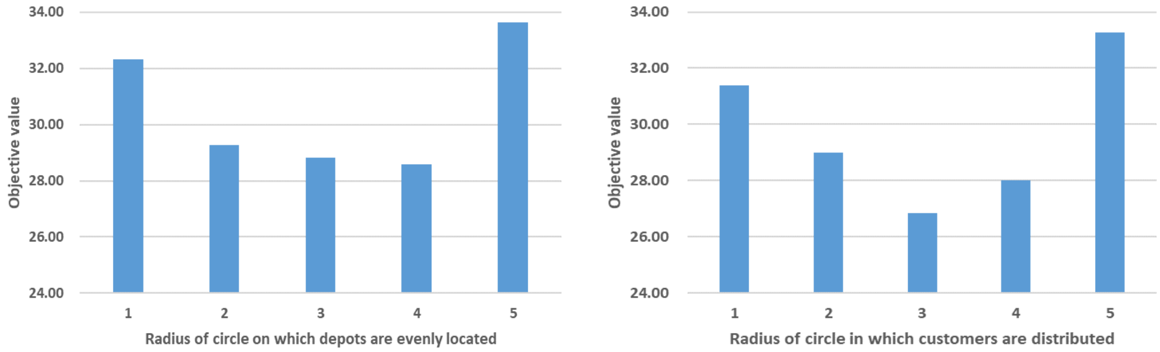


Figure 6: Sensitivity analysis of regions of the customers and the depots

(4) Sensitivity analysis of distance between regions of customers and depots

The fourth series of experiments is a sensitivity analysis of the distance between the previous mentioned two regions assuming that the two circles are centred at different locations. The results obtained are shown in Figure 7. More specifically, the x-axis denotes the distance between the center of the circle of radius r_2 on which the three depots are located and the center of the circle of radius r_1 in which 24 customers and 5 external delivery tasks are distributed. In the experiments, the first circle's radius is four and the second circle's radius is five. The blue bars in Figure 7 demonstrates the increasing trend of the final cost for increasing distances. The orange curve in Figure 7 shows that the instances are more difficult to solve, as testified by the increase in the computation time of method CG.

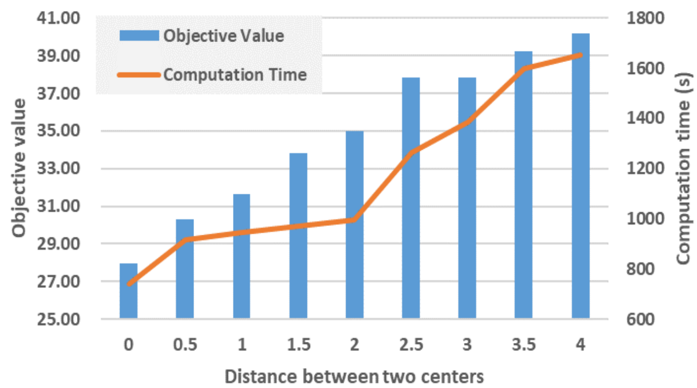
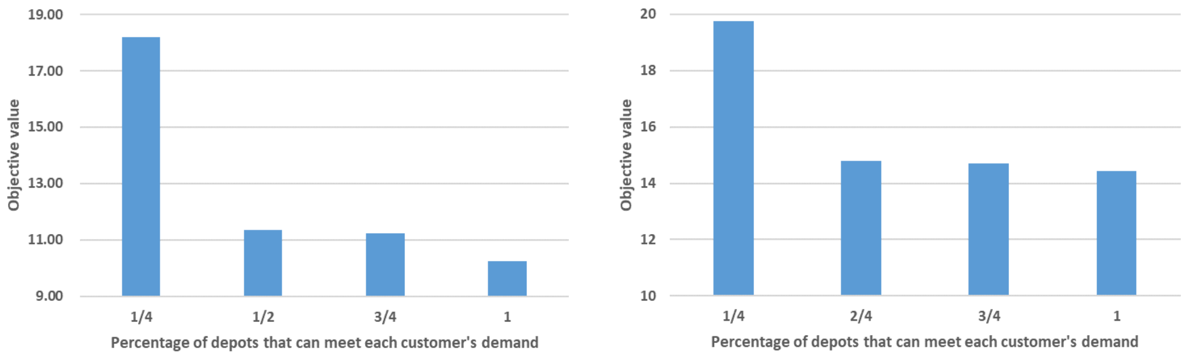


Figure 7: Sensitivity analysis of distance between regions of customers and depots

(5) Sensitivity analysis of the homogenous degree of depots

The last series of experiments is a sensitivity analysis of the percentage of depots that can meet the customer's demands. Figure 8 illustrates the results for the two series of experiments, both of which are based on the instances with four depots. More precisely, we consider different configurations of the depots, and we compute for each customer the number of depots that can be used to serve the customer, and the corresponding average value. In the figure, values "1/4, 2/4, 3/4, and 1" on the x-axis denote the different depot configurations, where value x/4 means that on average x depots can be used (over the four depots) to serve a customer. The case with the percentage "1" means that the depots are *homogenous* or identical. As expected, the results show that the cost decreases for increasing values of degree of homogeneity of the depots. Indeed, when a logistics company (or retailer) deploys comprehensive goods in all the depots or even makes all the depots storing all the types of goods, there are benefits in term of cost reduction. On the other side, this can increase the running cost associated with the depots.

Figure 8 also shows that a trade-off exists between the degree of homogeneity of the depots and the last-mile delivery activities.



4 depots, 6 and 2 dedicated and occasional vehicles, 18 customers, 4 external delivery tasks

4 depots, 6 and 5 dedicated and occasional vehicles, 24 customers, 5 external delivery tasks

Figure 8: Sensitivity analysis of percentage of depots that can meet each customer's demand

6. Solving large-scale real-world instances

In this section, we show the effectiveness of the CG method in solving large-scale real-world instances provided by a prominent e-commerce logistics platform, JD Logistics. The data of the instances come from China's largest district, Pudong district in Shanghai city, which has a territory of about 1210 km^2 and a population of 5.5 million. In this district, JD Logistics has established 20 primary depots (also called *stations*). Figure 9 depicts the locations of the depots, which are marked with black squares in the figure. The real instances considered in our experiments involve 200 internal deliveries, 40 external deliveries, and 30 occasional vehicles. Figure 9 also depicts the different locations; the data of the instance are summarized in Table 7.

Table 7: Instance data for the real case of JD Logistics in Shanghai Pudong district

	depot numbers $\|D\|$	dedicated vehicle numbers $\|K\|$	occasional vehicle numbers $\|K'\|$	total vehicle numbers $\|K\|+\|K'\|$	customer points $\|N\|$	external delivery tasks $\|L\|$
ISG12	20	40	30	70	200	40

It is time-consuming to solve the pricing problem for this kind of large-sized instance using a general-purpose solver like CPLEX (as in our previous experiments). Hence, we designed a new, more efficient algorithm to solve the pricing problem of the CG method based on a Particle Swarm Optimization (PSO) algorithm (Sanchez et al., 2017). PSO has proven to be very effective in solving hard combinatorial optimization problems such as this. To conserve space, we do not report the PSO used in our experiments in detail. Clearly, the PSO algorithm does not guarantee the optimality of the solutions of the pricing problem, and therefore the resulting solution cost of LR-RMP is no more than a valid lower bound on the optimal solution cost of the problem.

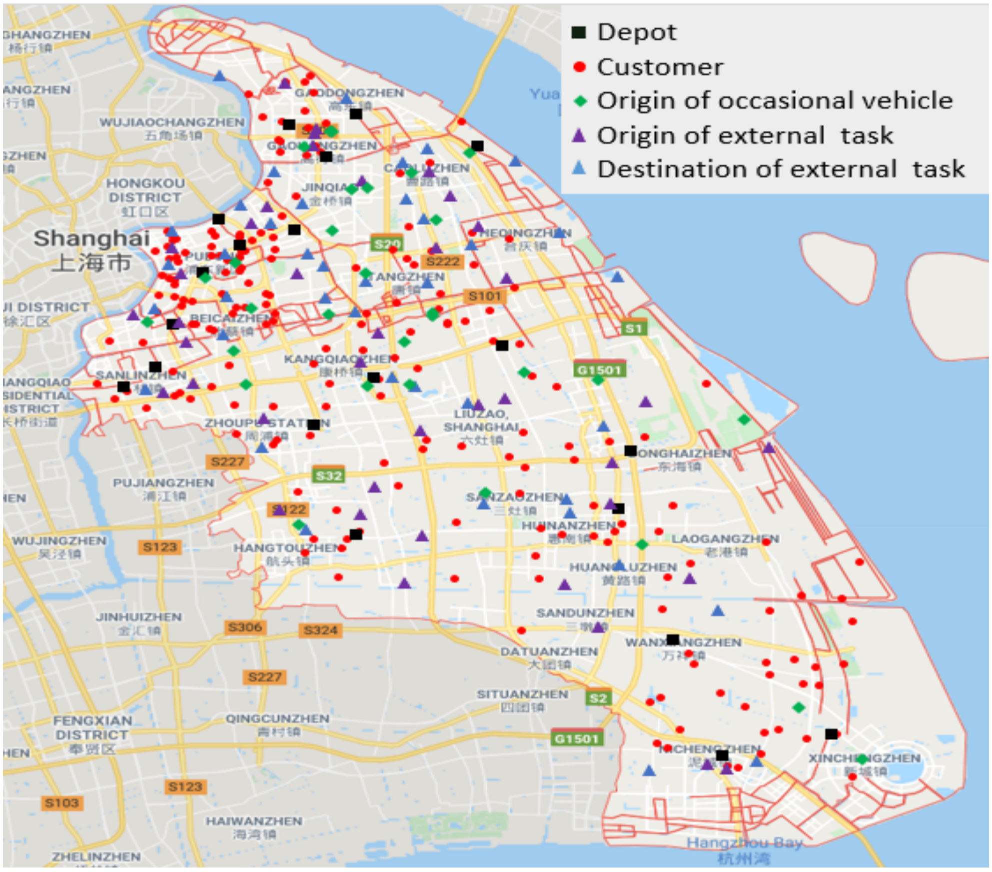


Figure 9: Layout of the real instance of JD Logistics in Shanghai Pudong district

Table 8 summarizes the results obtained using the CG method with the PSO algorithm for the pricing problem (denoted as CG-PSO) for the real instances of JD Logistics.

Table 8: Results of method CG-PSO on real-world instances of JD Logistics

Instances		Model		Rule	Gap
scale	index	F_{CG}	$t_{CG}(s)$	F_{DR}	
ISG12	1	40.83	1733	130.08	218.59%
	2	44.89	7015	144.55	222.01%
	3	57.12	2870	160.41	180.83%
	4	44.80	2229	129.36	188.75%
	5	56.51	2526	142.72	152.56%
Average					192.55%

In Table 8, we compare CG-PSO with the rule-based approach introduced in the previous sections, as a way of comparing a qualitative approach versus the quantitative approach provided by the CG method. The results show that CG-PSO can bring significant savings, as testified by the average gap of 193%. Given the difficulty of the instances, the computing time of method CG-PSO increases, but it remains acceptable and thus is usable in practice.

7. Conclusions and future research

In this paper, we investigate a vehicle routing problem that arises in the e-commerce context when a mixed fleet of dedicated and occasional vehicles must be routed to serve a set of heterogeneous orders composed of internal and external deliveries. We describe new mixed-integer programming models for the problem together with a column generation–based heuristic, denoted as CG. The models and CG-based method were validated through an extensive computational study that also included results for real-world instances. The study makes four main contributions:

1) The problem addressed in this paper constitutes a new variant with respect to the existing literature on vehicle routing problems, which is more realistic in the context of the e-commerce logistics industry adopting the crowdsourced delivery model and serving both internal and external deliveries.
2) The numerical experiments showed that the proposed CG method can reach, on average, an optimality gap of 0.11% in a limited amount of computing time. The effectiveness of the CG method is also validated by significant cost savings in comparison with two common intuitive decision rules.
3) The sensitivity analysis reported in the paper indicates useful managerial implications. In particular, the analysis shows that trade-offs exist in the composition of the vehicle fleet and in the definition of the degree of homogeneity of the depots and last-mile delivery activities.
4) The CG method was also used to solve a large-scale real-world problem provided by a leading e-commerce logistics platform (JD Logistics) involving 20 depots, 200 internal customers, 40 external tasks, and 70 vehicles. The results obtained shows that the CG method brings significant benefits in comparison with intuitive decision rules. Furthermore, it produces high-quality solutions in less than one hour of computing time, a level of computational efficiency that means it is of practical use.

We did not consider the behavioral factors of occasional drivers or pricing decisions, topics that have been considered by some studies of shared mobility studies (Qi et al., 2018). Further, the uncertainty of the actual travelling time between each link is also important in practice (Laporte et al., 1992). These topics, together with extensions of our algorithms to a larger scope of applications, will be the subject of our future work.

References

Alinaghian M., Shokouhi N. (2018) Multi-depot multi-compartment vehicle routing problem, solved by a hybrid adaptive large neighborhood search. *Omega* 76: 85−99.

Alnaggar A., Fatma G., Bookbinder J. H. (2019) Crowdsourced delivery: a review of platforms and academic literature. *Omega*, in press.

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.

Azadeh A., Farrokhi-Asl H. (2019) The close-open mixed multi depot vehicle routing problem considering internal and external fleet of vehicles. *Transportation Letters* 11(2): 78−92.

Baldacci R., Mingozzi A, Roberti. R. (2011) New route relaxation and pricing strategies for the vehicle routing problem. *Operations Research* 59 (5): 1269–83.

Baldacci R., Mingozzi A, Roberti. R. (2012) Recent exact algorithms for solving the vehicle routing problem under capacity and time window constraints. *European Journal of Operational Research* 218 (1): 1–6.

Baldacci R., Mingozzi A., Roberti R., Calvo R.W. (2013) An exact algorithm for the two-echelon capacitated vehicle routing problem. *Operations Research* 61(2): 298−314.

Bent R., Hentenryck P. V. (2004) A two-stage hybrid local search for the vehicle routing problem with time windows. *Transportation Science* 38(4): 515−530.

Bergmann F.M., Wagner S.M., Winkenbach M. (2020) Integrating first-mile pickup and last-mile delivery on shared vehicle routes for efficient urban e-commerce distribution. *Transportation Research Part B* 131: 26−62

Bertsimas D., Jaillet P., Martin S. (2019) Online vehicle routing: the edge of optimization in large-scale applications. *Operations Research* 67(1):143−162.

Bianchessi N., Drexl M., Irnich S. (2019) The split delivery vehicle routing problem with time windows and customer inconvenience constraints. *Transportation Science* 53(4): 1067−1084.

Cuda R., Guastaroba G., Speranza M.G. (2015) A survey on two-echelon routing problems. *Computers & Operations Research* 55: 185–99.

Dantzig G. B., Ramser J. H. (1959) The truck dispatching problem. *Management Science* 6(1): 80−91.

Dantzig, G. B., Wolfe, P. (1960) Decomposition principle for linear programs. *Operations Research*, 8:101–111.

Desrochers M., Desrosiers J., Solomon M. M. (1992) A new optimization algorithm for the vehicle routing problem with time windows. *Operations Research* 40(2), 342−354.

Dimitrakos T. D., Kyriakidis E. G. (2015) A single vehicle routing problem with pickups and deliveries, continuous random demands and predefined customer order. *European Journal of Operational Research* 244(3): 990−993.

Francois V., Arda Y., Crama Y. (2019) Adaptive large neighborhood search for multitrip vehicle routing with time windows. *Transportation Science* 53(6): 1706−1730.

Gansterer M., Hartl R. F. (2018) Collaborative vehicle routing: a survey. *European Journal of Operational Research* 268(1): 1−12.

Lahyani R., Gouguenheim A.-L., Coelho L.C. (2019) A hybrid adaptive large neighbourhood search for multi-depot open vehicle routing problems. *International Journal of Production Research* 57(22): 6963−6976.

Lalla-Ruiz E., Exposito-Izquierdo C., Taheripour S., Voss S. (2016) An improved formulation for the multi-depot open vehicle routing problem. *OR Spectrum* 38(1): 175−187.

Laporte G., Nobert Y., Taillefer S. (1988) Solving a family of multi-depot vehicle routing and location-routing problems. *Transportation Science* 22(3): 161−172.

Laporte G., Louveaux F., Mercure H. (1992) The vehicle routing problem with stochastic travel times. *Transportation Science* 26(3): 161−170.

Laporte G. (2009) Fifty years of vehicle routing. *Transportation Science* 43(4): 408−416.

Liu R., Jiang Z. (2012) The close–open mixed vehicle routing problem. *European Journal of Operational Research* 220(2): 349−360.

Liu R., Jiang Z., Geng N. (2014) A hybrid genetic algorithm for the multi-depot open vehicle routing problem. *OR Spectrum* 36(2): 401−421.

Lübbecke, M. E., Desrosiers, J. (2005) Selected topics in column generation. *Operations Research*, 3:1007–1023.

Mahmoudi M., Zhou X. (2016) Finding optimal solutions for vehicle routing problem with pickup and delivery services with time windows: A dynamic programming approach based on state-space-time network representations. *Transportation Research Part B* 89: 19−42.

Montoya-Torres J. R., López Franco, J., Nieto Isaza, S., Felizzola Jiménez, H. and Herazo-Padilla, N. A literature review on the vehicle routing problem with multiple depots. *Computers & Industrial Engineering* 79: 115–29.

Nagy G., Wassan N. A., Speranza M. G., Archetti C. (2015) The vehicle routing problem with divisible deliveries and pickups. *Transportation Science* 49(2): 271−294.

Paradiso R., Roberti R., Lagana D., Dullaert W. (2020) An exact solution framework for multitrip vehicle-routing problems with time windows. *Operations Research* 68(1): 180−198.

Pecin D., Contardo C., Desaulniers G., Uchoa E. (2017) New enhancements for the exact solution of the vehicle routing problem with time windows. *INFORMS Journal on Computing* 29(3): 489−502.

Polat O., Kalayci C.B., Kulak O., Günther H.-O. (2015) A perturbation based variable neighborhood search heuristic for solving the vehicle routing problem with simultaneous pickup and delivery with time limit. *European Journal of Operational Research* 242(2): 369−382.

Qi W., Li L., Liu S., Shen Z.-J. M. (2018) Shared mobility for last-mile delivery: design, operational prescriptions, and environmental impact. *Manufacturing & Service Operations Management* 20(4): 601−800.

Salhi S., Imran A., Wassan N.A. (2013) The multi-depot vehicle routing problem with heterogeneous vehicle fleet: formulation and a variable neighborhood search implementation. *Computers & Operations Research* 52: 315−325.

Sanchez, I.A.L., Vargas, J.M., Santos, C.A., Mendoza, M.G., Moctezuma, C.J.M. (2017) Solving binary cutting stock with matheuristics using particle swarm optimization and simulated annealing. *Soft Computing* 22(1): 1-9.

Toth P., Vigo D. (2014) *Vehicle Routing: Problems, Methods, and Applications, Second Edition*. Society for Industrial and Applied Mathematics.

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.

Wang K. Zhen L. Wang S., Laporte G. (2018) Column generation for integrated berth allocation, quay crane and yard assignment problem. *Transportation Science* 52(4): 812−834.

Yang H., Qin X., Ke J., Ye J. (2020) Optimizing matching time interval and matching radius in on-demand ride-sourcing markets. *Transportation Research Part B* 131: 84−105.

Yao Y., Zhu X., Dong H. Wu S., Wu H., Tong L.C., Zhou X. (2019) ADMM-based problem decomposition scheme for vehicle routing problem with time windows. *Transportation Researach Part B* 129: 156−174.