CIRRELT-2021-46

Crowdshipping: An Open VRP Variant with Stochastic Destinations

Fabian Torres
Michel Gendreau
Walter Rei

November 2021

Bureau de Montréal
Université de Montréal
C.P. 6 128, succ. Centre-Ville
Montréal (Québec) H3C 3J7
Tél : 1-514-343-7575
Télécopie : 1-514-343-7121

Bureau de Québec
Université Laval,
2325, rue de la Terrasse
Pavillon Palasis-Prince, local 2415
Québec (Québec) G1V 0A6
Tél : 1-418-656-2073
Télécopie : 1-418-656-2624

Crowdshipping: An Open VRP Variant with Stochastic Destinations

Fabian Torres[1,2,*], Michel Gendreau[1,2], Walter Rei[1,3]

[1] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)
[2] Department of Mathematics and Industrial Engineering, Polytechnique Montréal
[3] Department of Management and Technology, Université du Québec à Montréal

Abstract. E-commerce continues to grow throughout the world due to people's preference to stay at home rather than going to a brick-and-mortar retail store. COVID-19 has exacerbated this trend. Concurrently, crowd-shipping has been gaining in popularity due to both the increase in e-commerce and the current pressures due to COVID-19. We consider a setting where a crowd-shipping platform can fulfill heterogeneous delivery requests from a central depot with a fleet of professional vehicles and a pool of capacitated occasional drivers. We divide delivery requests into sectors to represent different neighborhoods in a city. Occasional drivers have unknown destinations that can be any-where inside the sectors. Route duration constraints are modeled to motivate participation and increase the probability of route-acceptance by keeping routes short. We assume that occasional drivers will choose routes that are better compensated and that the probability of route-acceptance is dependent on other routes being offered. We propose a two-stage stochastic model to formulate the problem. We use a branch-and-price algorithm capable of solving 50-customer instances, and develop a heuristic that can solve larger 100-customer instances quickly. An upper bound for the total number of occasional drivers is used to reduce the number of constraints in the master problem and reduce the complexity of the pricing problems. We show that occasional drivers with destinations far from the depot reduce the cost by over 30%, while occasional drivers with destinations that are near the depot reduce the cost by 20%. We show that route duration constraints and capacity constraints can restrict the occasional driver routes and both need to simultaneously increase in order to have cost reductions. This setting of crowd-shipping is a viable option for last-mile deliveries.

Keywords: Crowd-shipping, crowd-logistics, crowd drivers, occasional drivers, city logistics, stochastic programming, dynamic programming, column generation.

Acknowledgements. This research was partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) under its Discovery Grants research program, as well as the Canada Research Chair in Stochastic Optimization of Transport and Logistics Systems by Polytechnique Montréal, and by the Ecuadorian Secretaría de Educación Superior, Ciencia,Tecnología e Inovación (SENESCYT) under its international scholarships program.

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

---

* Corresponding author: fabiantodu@gmail.com

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2021

© Torres, Gendreau, Rei and CIRRELT, 2021

1 Introduction

E-commerce continues to grow throughout the world due to people's preference to stay at home rather than going to a brick-and-mortar retail store. COVID-19 and the increasing risk of global pandemics has exacerbated this tendency Bhatti et al. (2020); Gao et al. (2020). People would rather stay at home to minimize the risk of infection. At the same time, jobs are being lost at previously busy restaurants and public stores Sanchez et al. (2020). The jobs that are lost, according to Sanchez et al. (2020), tend to be low paid and less secure, and are usually held by young, poorly educated workers and migrants. Fortunately and concurrently, crowd-shipping has been gaining in popularity due to e-commerce and COVID-19. Platforms like Amazon Flex and Uber-Eats have increasing market participation. Raj et al. (2020) show that restaurants have increased sales through Uber-Eats, a food delivery service provided by Uber. Crowd-shipping has the potential to alleviate the unstable predicament of workers by providing a means to add an additional income. The idea of crowd-shipping is to utilize the vehicles that are already on the road (e.g., personal vehicles), to deliver packages. In a recent survey by Sampaio et al. (2019), the challenges and opportunities of crowd-shipping are emphasized. Retailers, e.g., Amazon, have traditionally owned their fleet of professional vehicles (PV) to perform delivery requests. With crowd-shipping they can now depend as well on a fleet of crowd-drivers or occasional drivers (OD) that can fulfill some delivery requests. ODs are individuals with a personal vehicle that have a planned trip with a destination and are willing to perform delivery tasks.

The challenge is the unknown preferences and characteristics of ODs. For instance, ODs have different models of vehicles of heterogeneous sizes, thus, the capacity of OD-vehicles varies and the retailer would not know how many packages an OD can deliver *a priori*. In addition, ODs can accept or reject routes. Thus, the probability that a route will be accepted by ODs is dependent on other routes that could be less or more attractive to ODs. OD-destinations can be different and change from day to day, and overall OD-participation from day to day is uncertain.

The most successful crowd-shipping platform (CSP) is Amazon Flex that now operates in the US and Canada, and in over 100 cities AmazonFlex (2021). Amazon has solved some of the challenges by creating a program where individuals can sign up *a priori* only if they meet some basic requirements. Participants must pass security checks and complete basic workshops to qualify. Walmart recently started its own CSP called Walmart Spark Delivery. For a review of the different CSPs and the scientific literature on crowd-shipping, see Alnaggar et al. (2021). They noticed that most of the real-world CSP have capacitated crowd-drivers that perform deliveries from a central depot, while most of the scientific literature on crowd-shipping considers that ODs can perform only one delivery regardless of the capacity and considers multiple pickup locations. Few stochastic variants are identified.

In this paper, we consider a setting where a CSP has a set of delivery requests that can be fulfilled by a fleet of PVs or a pool of capacitated ODs who have previously registered to the platform. Deliveries must be fulfilled from a central depot. The products sold are heterogeneous, e.g., microwave ovens, books, pencils, frozen foods. Some items have to be delivered within a time window. ODs are capacitated and have different random destinations that change from day to day. ODs can choose routes, thus, the probability that ODs accept a route is dependent on the other routes that are offered to them. Additionally, cities are generally divided by neighborhoods and natural geographic barriers, e.g., rivers, mountains, road networks and parks. In practice, customers are generally clustered in different areas or sectors in a city. Thus, we assume that delivery requests can be clustered into sectors and that OD-destinations are in a city sector. Furthermore, ODs do not want to deviate too far from their trajectory.

We assume that historical information exits on OD-destinations for each sector, which allows the creation of a discrete probability function that can predict the supply of ODs with destinations that end in the sector. The setting is suitable for large data-rich organizations that can gather such information over the course of their operations, e.g., Amazon. In this manner, each sector has a discrete probability function that predicts the number of ODs that will end their trip in the sector. In addition, the probability of route rejection is dependent on the other routes offered to ODs.

The main contributions of this paper are the following:

- We introduce a new variant for crowd-shipping: Open VRP with stochastic destinations.
- In order to increase participation, we propose a route duration constraint for OD-routes, so that the routes built for ODs do not exceed a certain length. In this manner, ODs that participate in the CSP know in advance that their route will not exceed a certain value.
- We extend the set covering formulation presented by Torres et al. (2020), by adapting the optimization model for this new variant.
- We strengthen the upper bounds on the total number of ODs used in a solution proposed by Torres et al. (2020). Specifically, we use the route duration constraint and information about the shortest route to strengthen the upper bound.
- We develop a column generation heuristic to provide solutions to larger instances.
- Finally, we provide extensive computational experiments that show valuable insights to the VRP with stochastic destinations.

The remainder of this paper is organized as follows: In Section 2, we review the related literature. In Section 3, we describe the problem and we present a set covering formulation. In Section 4, we describe the solution approach utilized to solve the problem. In Section 5, we perform extensive computational experiments and provide valuable insights about the characteristics of the problem. Finally, in Section 6, we conclude and provide interesting future research directions.

2 Related work

After crowd-shipping was conceptualized by Amazon and other companies, an initial quantitative study of crowd-shipping was presented by Archetti et al. (2016). The authors introduce the Vehicle Routing Problem with Occasional Drivers (VRPOD). They formulate a deterministic and static model where a set of delivery requests have to be fulfilled from a central depot by a mixed fleet of unlimited PVs that complete closed routes, and a set of ODs that are willing to deliver some packages while they travel towards their specific destination. The compensation of in-store customers could be based on different strategies. Two different compensation schemes are considered, the first independent and the second dependent of the OD-destination. An mixed integer program (MIP) is introduced to solve small instances, and a multi-start heuristic is proposed to provide solutions to larger instances. It is shown that significant cost reductions are achieved by employing the static-deterministic ODs compared to conventional vehicle routing delivery plans.

A main question that arises from Archetti et al. (2016) is how can ODs be applied in a dynamic and/or stochastic setting? In practice, the flow of information is going to occur in stages. ODs could arrive dynamically throughout the day; the total number of ODs is unknown until late in the day. Furthermore, OD could reject or accept routes. In Gdowska et al. (2018), the authors consider the possible rejection of delivery tasks by ODs. The probability of rejecting a delivery request by an OD is viewed as independent from other delivery requests. They extend the MIP presented in Archetti et al. (2016) to a stochastic setting by determining beforehand the set of customers that will be proposed to ODs based on an expected cost function. Initially, the set of delivery requests is separated into two sets by a heuristic method; the first set of delivery requests is fulfilled by PVs and the second set is satisfied by ODs. However, ODs can decline routes, resulting in a penalty per delivery request contained in the OD-set and not fulfilled due to OD rejection. Recall that ODs in this setting can only visit one customer, therefore, OD routes contain a single customer. Every delivery request has a probability of being accepted by an OD. Thus, the expected cost function determines the expected cost of the OD-customer set based on a fixed compensation and the expected penalty if a customer is not served. The remaining PV-customer set is solved optimally with the model proposed by Archetti et al. (2016) and a commercial solver. Due to the complexity of the VRP, only small instances with 15 delivery requests were solved. Even though the resulting VRP is solved optimally, it is important to note that this method is not exact. However, an important contribution of this study was to present the idea of separating the set of delivery requests into two sets and accounting for the possibility of route rejection. This can be useful in practice since PVs can start some long routes at the beginning of the planning horizon before knowing the number of ODs that will be available or that will not reject routes.

In a similar work, Torres et al. (2020) presented a two-stage stochastic framework that we extend in this paper. They consider a setting where a crowd-shipping platform delivers heterogeneous packages from a central depot. A fleet of unlimited PVs and a pool of ODs are used to fulfill the delivery requests. The ODs are capacitated and can deliver multiple packages, hence planning routes for ODs is required. Failed delivery attempts are possible in the setting, and thus, ODs must return to the depot if they have undelivered packages. Hence, ODs do not have a fixed destination and are willing to return to the depot. Similar to Gdowska et al. (2018), in the first-stage delivery requests are separated into two sets: one set of delivery requests is assigned to the PV-fleet, and the other is assigned to be visited by ODs. Route rejection is considered by a discrete probability function that predicts the participation of ODs or acceptance of OD routes. However, different from Gdowska et al. (2018), the set separation step is done through the following two-stage stochastic set-partitioning formulation of the VRP:

$$\min \sum_{r \in \Omega \cup \Omega'} c^r \lambda^r + \mathscr{Q}(\lambda) \tag{1}$$

$$\text{s.t.} \quad \sum_{r \in \Omega \cup \Omega'} a_i^r \lambda^r = 1 \qquad \forall i \in N \tag{2}$$

$$\sum_{r \in \Omega'} \lambda_r \leq M \tag{3}$$

$$\lambda_r \in \{0, 1\} \qquad \forall r \in \Omega \cup \Omega'$$

The set Ω is the set of all feasible routes for PVs and Ω' is the set of all feasible routes for ODs. Both sets contain an exponential number of routes. Binary variables λ^r denote if route $r \in \Omega \cup \Omega'$ is used in a solution. Parameters c^r are the cost of route $r \in \Omega \cup \Omega'$ and parameter M is an upper bound on the total number of ODs that can be used, e.g, the total number of ODs registered in the platform. The parameter a_i^r is equal to one if and only if customer $i \in N$ is visited in route r, where N is the set of customers. The first-stage solution gives a natural separation of the delivery requests that are to be fulfilled by PVs and ODs based on routes. The second-stage cost, i.e., $\mathscr{Q}(\lambda)$, is the expected cost incurred by applying recourse actions and penalties for the first-stage subset of delivery requests that is assigned to ODs. The probability of a route being rejected is dependent on other routes. ODs might reject routes or simply not participate; in which case, the recourse actions consist in having larger PVs complete the OD-routes that are not fulfilled. The PVs that must be used in the recourse action to fulfill OD-routes are considered to be more expensive, i.e., a parameter $\alpha > 1$ multiplies the regular cost of a PV.

The model is then linearized by adding more binary variables to the model. A branch-and-price algorithm is used to solve some large 100-customer instances; a column generation heuristic is also shown to perform faster. Computational experiments show the value of the stochastic solution for this model to be over 20% for some cases. However, this study does not address the problem when ODs have planned trips with a desired destination. In some settings, requiring ODs to return to the depot can be limiting and/or increase the cost of the operation. In this paper, we extend the framework of Torres et al. (2020) to a case where ODs have unknown destinations. The resulting problem is an Open VRP variant with stochastic destinations.

Dayarian and Savelsbergh (2017) introduce a highly dynamic, same-day delivery environment, where both ODs and delivery requests arrive dynamically throughout the planning horizon. A program is created that requires ODs to describe some of their characteristics with the intention of reducing the uncertainty. More specifically, the willingness of ODs to deliver to online customers is represented by a coverage area that describes the location of customers that will be accepted by the OD. It is assumed that ODs will accept delivery tasks as long as all delivery locations are inside the coverage area and that the regular fleet of vehicles has an unlimited capacity. A sample scenario planning approach is developed and the routing problem is solved using a tabu search heuristic.

When ODs have a set of routes to choose from, the probability that a route will be accepted or rejected depends on the other routes available. Some routes might be more attractive for ODs that others.

Table 1 shows the summary of the stochastic variants that were identified in the literature.

Study	**OD-Destination**	**OD-Capacity**	**Probability**	**Recourse**
Dayarian and Savelsbergh (2017)	YES	NO	Independent	NO
Gdowska et al. (2018)	YES	NO	Independent	NO
Torres et al. (2020)	NO	YES	Dependent	YES
Dahle et al. (2017)	YES	NO	Independent	NO
This paper	YES	YES	Dependent	YES

Table 1: Summary of stochastic variants

As noted by Alnaggar et al. (2021), the capacity of ODs is rarely considered in the literature except by a few papers, e.g., Archetti et al. (2021). In addition, stochastic variants in VRPs generally present a recourse action in the case the routing plan fails, e.g., see Gendreau et al. (2014). In practice, when an unexpected event occurs an action needs to be taken to mitigate the losses. To the best of our knowledge, only one stochastic variant proposed a recourse action for ODs, i.e., Torres et al. (2020).

For additional interesting dynamic variants, see Dahle et al. (2017), Arslan et al. (2019) and Archetti et al. (2021). For static-deterministic variants, see Macrina et al. (2020), Macrina and Guerriero (2018) and Dahle et al. (2019).

3 Problem Description and Formulation

In this section, we describe and formulate the problem. In Section 3.1, we describe the way we group customers into sectors and describe the characteristics of each sector. In Section 3.2, we delineate the routes that will be offered to ODs. In Section 3.3, we describe the CSP, the flow of information and the cost structure. Finally, in Section 3.4, we formulate the problem.

3.1 Sectors

Cities are generally divided by natural barriers, e.g., rivers, mountains, or by large parks; cities are also divided by industrial areas, neighborhoods and the road network. Thus, we assume that a set of delivery requests can be easily clustered to represent different sectors within a city. In figure 1, we can see a description of an instance of the problem.

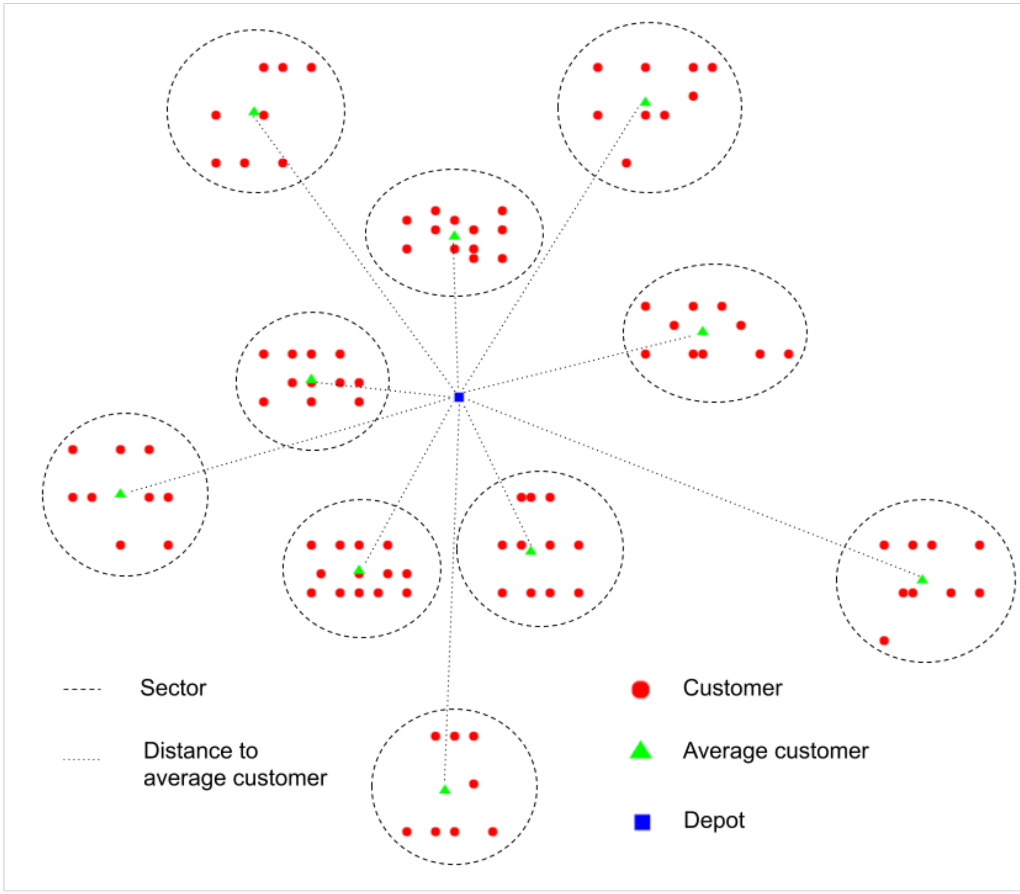


Figure 1: Example of instance with stochastic destinations

- **Customers**. The set of customers in a sector represents delivery requests that must be fulfilled by the platform on a given day. Each customer has a set of coordinates describing her location, a demand to represent the volume of the parcel, and an earliest and latest arrival time, i.e., a time window.

- **Average customer**. The average customer is the average location of all delivery requests in the sector. This average customer represents the center of a sector in a city so that all the other customers in the sector are at a reasonable distance from this center. It facilitates estimations of route duration constraints as well as the compensation.

- **Distance to sector**. Distance from the depot to the average customer, i.e., D_u for each sector $u \in U$, where U is the set of sectors.

3.2 OD-routes

A description of the routes that will be offered to ODs is presented in this section. Figure 2 shows an example of 3 OD-routes.

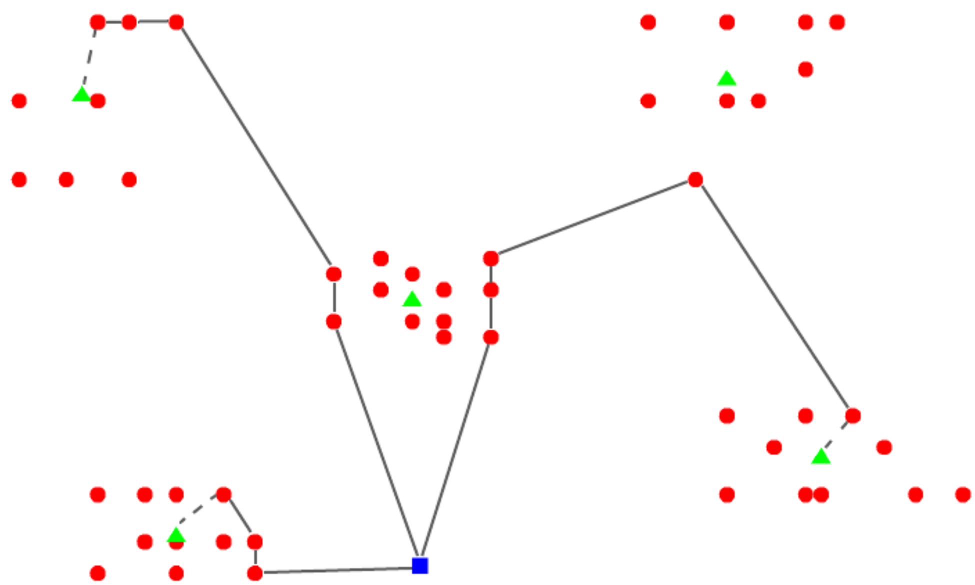


Figure 2: Example of OD-routes

3.2.1 Total load

We assume that OD-vehicles have a capacity of at least Q'. While ODs may have vehicles with a larger capacity, the total load of OD-routes cannot exceed the capacity Q'. In this way, all routes created are feasible for ODs to complete. However, they can be rejected by ODs, see section 3.2.3. The fleet of PVs is assumed to have a PV-capacity of Q, that is strictly larger than Q'.

3.2.2 Route duration constraint

We assume that ODs are not willing to deviate too far from their trajectory as they go to their destination. Route duration constraints ensure that routes created for each sector are not too long, thus, increasing participation and route acceptance. The total distance traveled in the route from the depot to the average customer in the sector cannot exceed a value of $\bar{D}_u$ for $u \in U$. The value $\bar{D}_u$ is the route duration for the sector that contains the planned destination of the OD and it is defined as follows:

$$\bar{D}_u = \phi D_u + \tau. \tag{4}$$

Parameter ϕ is the acceptable deviation from the OD-trajectory as an additional percentage of D_u for each sector $u \in U$. Parameter τ is an additional value added to D_u. The two parameters, i.e., ϕ and τ, give the platform flexibility to create routes that can be acceptable to ODs that have a destination near or far from the depot. Some ODs that have a destination in close proximity to the depot and have extra time, might prefer a higher value of τ to increase their income.

3.2.3 Probability of OD-route acceptance

In a recent study Asdecker and Zirkelbach (2020), it is shown that crowd delivery drivers are primarily motivated by compensation. In addition, Torres et al. (2020) show that, in their setting, there is no difference in assigning the most expensive routes to ODs first or expecting ODs to choose these routes based on compensation given the cost structure proposed in their model. Thus, in this paper, we assume that all ODs prefer higher compensated routes and we assume that higher compensated routes are also more expensive for the CSP. The probability that a route has of being accepted by the pool of ODs depends on the other routes that are offered to ODs. For instance, let $r_1, \ldots, r_s$ be OD-routes listed by higher to lower compensation. The first ODs to arrive will prefer route r_1 since it has the highest compensation. The second ODs will not be able to choose route r_1 and will have to choose r_2, i.e., the highest compensated route available. If we have any list of routes ordered from highest compensated to lowest, the highest compensated will have a higher probability of been chosen by ODs since it is their first choice. This implies that for any set of OD-routes given, we must order them first from highest to lowest compensated to know the probability of acceptance of each route. Each sector has a discrete probability distribution describing the probability of the supply of ODs, i.e., $P_u(\xi_u > s)$ where ξ_u is the expected supply of ODs that will have a destination in sector $u \in U$ and s is the OD-preference of the route w.r.t. the other routes in the sector. In this way, any list of OD-routes has to be ordered by compensation first and then we can calculate the probability of acceptance. This list of ordered preferences is commonly referred to as a preference ordering in economics Debreu (1954); Mrkela and Stanimirović (2021). Historical data on OD-destinations can be used to create the discrete probability function. ODs can reject routes or accept them; the historical data can show how many ODs have accepted routes on any given day for each sector.

3.2.4 Multiple sector visits

OD-routes can visit customers in multiple sectors $u \in U$ as long as the last customer of the route is in the sector of the OD-destination. This takes into account that ODs would prefer to be in close proximity to their final destination once they finish their routes. Figure 2 shows an example of 3 OD-routes. Two OD-routes visit customers that are in different sectors while the last one visits customers only in a single sector.

3.3 Crowd-shipping platform (CSP)

The CSP is a platform that requires ODs to register first in order to participate. If they meet certain requirements, e.g., security checks or attending workshops, then ODs can download an application and start delivering packages when they have time. These requirements can reduce uncertainty by standardizing the pool of ODs. Since the CSP creates the requirements, it can filter out ODs with small vehicles or ODs that are not interested in the type of delivery requests that they offer. Moreover, it allows ODs to understand how they will be compensated and what type of routes they are expected to perform. It is important to recall that, even if they are registered in the CSP, ODs are not forced to participate and can choose freely when they work and their destination from day to day. We assume ODs are rational agents that will choose the highest compensated routes, see section 3.2.3.

3.3.1 First stage

We consider a two-stage stochastic modelling framework. In the first stage, PV-routes and OD-routes have to be planned without exact knowledge of the supply of ODs that will arrive throughout the operations time span for each sector. PVs have to depart early to visit customers and perform long routes. This leaves a set of customers that need to be visited by ODs. The OD-routes are then offered to ODs in the CSP-app and they are expected to choose the highest compensated routes first, see section 3.2.3.

3.3.2 Second stage

The second stage begins at some point in time when the CSP cannot wait any longer for additional ODs to arrive to perform the remaining routes. At this point, recourse actions need to be taken to guarantee that the deliveries to all customers are made on time. Recourse actions consist of sending a PV to complete the OD-routes that were rejected by ODs.

3.3.3 Cost structure

Torres et al. (2020) consider a fixed compensation, i.e., F' and a variable compensation, i.e., $\beta' < 1$ for ODs. PVs also get paid a fixed cost, i.e., F and a variable cost that is set to 1. Variable costs are proportional to the distance traveled in the route. A cost per delivery is also considered for each route, i.e., w_r, for both PV and ODs. An additional recourse cost is incurred when OD-routes are not fulfilled by ODs and must be performed by PVs. The cost to use a PV to complete OD-routes is the regular cost of a PV times a penalty $\alpha > 1$. Let r^* be an OD-route that has been rejected and let d_{r^*} be the distance of this OD-route; the cost of a recourse action is $\alpha(F + d_{r^*} + w_{r^*})$. Thus, the additional cost incurred by performing a recourse action for the rejected OD-route r^* is the following:

$$\hat{z}_{r^*} = \alpha(F + d_{r^*} + w_{r^*}) - (F' + \beta' d_{r^*} + w_{r^*})$$

However, if a route "r" is not rejected by ODs the compensation paid to an OD is equal to: $\hat{c}^r = (F' + \beta' d_r + w_r)$. Here, both ODs and PVs are compensated to return to the depot if parcels are not delivered and a compensation per delivery is paid to both PVs and ODs. It should be noted that the compensation for the return trip is always paid, regardless of whether it is required or not. The expected cost of any given OD-route, including the cost associated with the risk of being rejected, is equal to the following:

$$\hat{c}_s^r = \hat{c}^r + \hat{z}_r P(\xi < s) \tag{5}$$

Here, the probability function (i.e., $P(\xi < s)$), returns the probability that a route will be rejected; ξ is the supply of ODs and the index s represents the priority, i.e., the order in which the CSP assigns routes to ODs based on the cost. As noted previously, there is little difference between expensive routes and better compensated routes.

In this paper, the priority is viewed as the preference of ODs (see Section 3.2.3). We consider the following additional cost incurred by a recourse action for a rejected OD-route r^*:

$$z_{r^*} = \alpha(F + d_{r^*} + d_{r^*}^0) - (F' + \beta' d_{r^*}) \tag{6}$$

A return trip to the depot is only expected for PVs, thus, the additional term $d_{r^*}^0$ is the distance from the last customer of the rejected route r^* back to the depot. However, we do not consider any compensation per delivery. The compensation given to an OD if the route is accepted is defined as $c_r = (F' + \beta' d_r)$; thus, the expected cost for any OD-route is:

$$c_{us}^r = c^r + z_r P_u(\xi_u < s) \quad \forall u \in U, \tag{7}$$

where $P_u(\xi_u < s)$ is the probability that route r will be rejected by ODs if it has a preference of "s" in the preference ordering of all OD-routes in sector $u \in U$. See Section 3.2.3 for an explanation of the probability function and the preference ordering of ODs.

3.4 Problem Formulation

We extend the model presented by Torres et al. (2020); where the model described in section 2, i.e.,(1)-(3), is linearized. Let λ_s^r be binary variables that are equal to one if route r is used in the solution with priority s. For simplicity, variables λ_0^r represent PV-routes.

$$\widehat{MP} = \min \sum_{r \in \Omega} \hat{c}_0^r \lambda_0^r + \sum_{s=1}^{M} \sum_{r \in \Omega'} \hat{c}_s^r \lambda_s^r \tag{8}$$

$$\text{s.t.} \quad \sum_{s=0}^{M} \sum_{r \in \Omega \cup \Omega'} a_{ir} \lambda_s^r \geq 1 \qquad \forall i \in N \tag{9}$$

$$\sum_{r \in \Omega'} \lambda_s^r \leq 1 \qquad \forall s \in \{1, \ldots, M\} \tag{10}$$

$$\lambda_s^r \in \{0, 1\} \qquad \forall s \in \{0, \ldots, M\}, r \in \Omega \cup \Omega'$$

The second-stage cost in model (1)-(3) is replaced by the sum of the binary variables λ_s^r times the cost (7). Constraints (9) make sure all delivery requests are covered by routes. Constraints (10) guarantee that there is only one route assigned to each priority. Interestingly, the order of routes from more expensive to cheapest is done automatically by the model; since this is a minimization problem, the model assigns routes to the best possible priority; for proof see Torres et al. (2020).

In this paper, we extend model (8)-(10); let λ_{us}^r be binary variables that are equal to one if route r is used in the solution with OD-preference s in sector $u \in U$. Then we define the master problem as follows:

$$MP = \min \sum_{r \in \Omega} c_0^r \lambda_0^r + \sum_{u \in U} \sum_{s=1}^{M_u} \sum_{r \in \Omega_l} c_{us}^r \lambda_{us}^r \tag{11}$$

$$\text{s.t.} \quad \sum_{u \in U} \sum_{s=0}^{M_u} \sum_{r \in \Omega \cup \Omega_u} a_{ir} \lambda_{us}^r \geq 1 \qquad \forall i \in N \tag{12}$$

$$\sum_{r \in \Omega_u} \lambda_{us}^r \leq 1 \qquad \forall s \in \{1, \ldots, M_u\}, u \in U \tag{13}$$

$$\lambda \in \{0, 1\}$$

The objective function (11) is to minimize the total cost of routing, including the expected recourse actions that will need to be taken. Set covering constraints (12) guarantee that all customers will be visited at least once. The set of constraints (13) allow only a single route to have an OD-preference of $s \in M_u$ per sector $u \in U$. The value M_u is an upper bound on the total number of OD-routes allowed in a sector; it can be equal to the total number of ODs registered to the CSP.

4 Solution Approach

Set covering formulations for VRP variants are usually solved by branch-and-price algorithms Poggi and Uchoa (2014); Costa et al. (2019). In this section, we present the approach used to solve our problem.

In Section 4.1, we propose a technique to determine the upper bound on the total number of ODs per sector in model (11)-(13), which is obtained by extending the one proposed by Torres et al. (2020). In Section 4.2 we describe the branch-and-price algorithm that is used to provide solutions to our problem. In Section 4.4, we present a column generation heuristic that is applied to quickly generate solutions.

4.1 Upper bound

Constraints (13) lead to a large number of constraints. In an effort to make the master problem smaller and more manageable, we derive an upper bound for M_u.

Let r be an OD-route with an expected cost c^r_{us} defined in equation (7), and let c^r_{u0} be the cost of deploying a PV to perform the OD-route in the first stage, i.e., without offering the route to ODs. If the expected cost for an OD is higher that the cost of employing a PV for the same route, i.e., if $c^r_{us} > c^r_{u0}$, then it should not be offered to ODs. The expected cost of an OD-route increases as the preference of the OD-route is lowered so that it is less likely to be accepted, i.e., $c^r_{us} < c^r_{us+1}$. A threshold probability of rejection for which an OD-route is less expensive for a PV is derived by Torres et al. (2020) as follows:

$$\hat{P} = \frac{\alpha - 1}{\alpha - \frac{F' + \beta' d_r}{F + d_r}} \tag{14}$$

Let the function $P^{-1}(.)$ be the quantile function of the discrete probability function of the supply of ODs. With the probability threshold in (14), we can estimate the maximum number of ODs in the model:

$$\hat{M} = P^{-1}\left(\frac{\alpha - 1}{\alpha - \frac{F' + \beta' d_r}{F + d_r}}\right)$$

By considering the limits of the distance traveled in route r, i.e., $d_r \to \infty$ and $d_r \to 0$, the following upper bound for the total number of ODs used in a solution is obtained:

$$\hat{M} = max\left[P^{-1}\left(\frac{\alpha - 1}{\alpha - \beta'}\right), P^{-1}\left(\frac{\alpha - 1}{\alpha - \frac{F'}{F}}\right)\right]$$

The first ratio is the threshold if the distance tends to infinity, i.e., $d_r \to \infty$. The second ratio gives the threshold if the distance tends to zero, i.e., $d_r \to 0$. The least restrictive of the two is used as an upper bound. In the problem setting that is considered in this paper, we have knowledge of the minimum and maximum distance of a route. The compensation structure is also slightly different. We adjust and improve the upper bound with this additional information. The route duration constraint value, i.e., $\bar{D}_u$, is the upper bound for the length of all OD-routes, and the distance from the depot to the closest customer in a sector, i.e., d_u, is the shortest distance for all OD-routes in sector $u \in U$. We improve the bound as follows:

$$\bar{M_u} = max\left[P_u^{-1}\left(\frac{\alpha - 1}{\alpha - \frac{F' + \beta' \bar{D}_u}{F + \bar{D}_u + d_u}}\right), P_u^{-1}\left(\frac{\alpha - 1}{\alpha - \frac{F' + \beta' d_u}{F + 2d_u}}\right)\right] \quad \forall u \in U \tag{15}$$

We calculate the threshold probability for each sector as shown in (15) the least restrictive threshold probability gives us an upper bound on the total number of ODs that are feasible in that sector, i.e., M_u.

4.2 Branch-and-Price (B&P)

The master problem (11)-(13) has an exponential number of variables. We consider a restricted version of the master problem called the restricted master problem that only has a few routes. We then generate the routes that are needed by solving pricing problems. For more information about B&P methods, see Wolsey (2020), Poggi and Uchoa (2014), Costa et al. (2019).

4.2.1 Pricing

The solution to the restricted master problem is not optimal since more columns could reduce the cost. Pricing is necessary to find all negative reduced cost columns, and if none can be found, to guarantee that columns with a negative reduced cost do not exist at the current iteration, and thus, proving that the solution for the restricted master problem is the same solution for the master problem. Let μ_{us} be the dual variables for constraints (13); let π_i be dual variables for constraints (12). In order to find the variable with the lowest reduced cost to add to the restricted master problem, the following problem has to be solved.

$$\min_{r,u,s} \hat{c} = \{c_{us}^r - \sum a_i^r \pi_i + \mu_{us} : r \in \Omega \cup \Omega', s \in \bar{M}_u, u \in U\} \tag{16}$$

Equation (16) describes a series of elementary shortest path problems with resource constraints (ESPPRC), which are NP-hard problems. The total number of ESPPRCs that need to be solved is equal to $\sum_u^U \bar{M}_u$. However, we can solve all of them in a single problem by using the cohesive pricing algorithm proposed by Torres et al. (2020) that can solve all problems in a single one. Route duration constraints are added as an additional resource in the algorithm and the destinations of ODs can be derived by the last extension of each route in the algorithm. Labeling algorithms have been used successfully for the ESPPRC with ng-routes and decremental state space relaxations (DSSR). The ng-routes relaxation consists of allowing non-necessarily elementary paths to be formed, i.e., customers can be visited more than once, as long as repeated customers are not contained on a specified set of nearest neighbors. In this way, only cycles that exit the neighborhood of nearest neighbors are allowed, these cycles tend to be more expensive since distant customers are visited,i.e., not nearest neighbors, and provide a better bound. DSSR gradually increases the customers in the ng-set of nearest neighbors from the empty set to a specified set cardinality, see Righini and Salani (2008). For further information about labeling algorithms and ng-routes relaxations, see Baldacci et al. (2011).

Let $\mathcal{L}_p = (i_p, l_p, t_p, d_p, \bar{\pi}_p, \mathscr{V}_p)$ be a label for a path p than ends at customer i_p, has a total load of l_p, a total time traveled of t_p, a total distance of d_p, cumulative dual variables equal to $\bar{\pi}_p$, and the set of unreachable customers equal to $\mathscr{V}_p$. We iterate over all possible loads l_p from 0 to the largest capacity, i.e., Q. Labels are extended from their current customer, i.e., i_p, to a reachable customer not contained in $\mathscr{V}_p$. Time window constraints for each customer must also be checked before an extension.

Dominance rules are then applied to eliminate labels that are not promising. For instance, let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two different labels with the same current customer, i.e., $i_1 = i_2$, and the same total load, i.e., $l_1 = l_2$. The following dominance rules are applied while $l_p \leq Q'$:

Dominance rules 4.1. *Label $\mathcal{L}_1$ dominates label $\mathcal{L}_2$ if the following rules are true:*

1. $t_1 \leq t_2$
2. $d_1 \leq d_2$
3. $\bar{\pi}_1 \geq \bar{\pi}_2$
4. $\mathscr{V}_1 \subseteq \mathscr{V}_2$

Once the total load of labels exceeds the OD-capacity, i.e., Q', then the paths described by the labels must be done by PVs. Route duration constraints are no longer necessary. The dominance rules applied once $l_p > Q'$ are the following:

Dominance rules 4.2. *Label $\mathcal{L}_1$ dominates label $\mathcal{L}_2$ if the following rules are true:*

1. $t_1 \leq t_2$
2. $d_1 - \bar{\pi}_1 \leq d_2 - \bar{\pi}_2$
3. $\mathscr{V}_1 \subseteq \mathscr{V}_2$

To add variables with the lowest reduced cost to the restricted master problem, we first look at labels that have a total load of less than or equal to the OD-capacity, i.e., $l_p \leq Q'$. The current customer of path p, i.e., i_p, belongs to a sector u_p, see Section 3.1. The paths that satisfy the route duration constraint for sector u_p are added to the restricted master problem only if they have a negative reduced cost. Paths that have a total load of more than the OD-capacity are added as PV-routes by extending back to the depot only if they have a negative reduced cost. Recall that the negative reduced cost of variables can be derived from (16).

4.2.2 Branching

We select the node of the tree with the best bound, then explore the problem. If the solution is fractional, we branch using the following order of importance:

1. Total number of vehicles, both PVs and ODs;
2. Total number of ODs;
3. Total number of ODs in each sector;
4. Finally, we branch on the flow between two customers.

4.3 Heuristic pricing

When solving the ESPPRC to find negative-reduced-cost columns, it is not necessary to find an optimal solution of the problem at each iteration; rather, any solution with a negative objective function value can be added to the restricted master problem. The difficulty of the exact pricing algorithm is caused by dominance rules 4.1.4 and 4.2.3. We drop these rules, and we maintain elementarity of all paths through feasibility checks. These two changes reduce the time complexity from exponential to pseudo-polynomial time. However, the guarantee of finding an exact solution is lost. Algorithm 1 is the procedure used to find columns and add them to the restricted master problem.

Algorithm 1: Heuristic pricing

```
j ←0;
found ← True;
while found do
    found ← False;
    Dominance rules(j), see Algorithm 2;
    Run labeling algorithm with feasible extensions;
    found ← Negative column found?;
    if found then
        Update the restricted master problem;
        j ← 0;
    else if j ≤ 1 then
        j ← j + 1;
        found ← True;
```

The procedure *Dominance rules*, described in Algorithm 2, sets the dominance rules to be used in Heuristic pricing. When $j = 0$, the total distance is sufficient to eliminate labels. This leads to a fast algorithm that runs in polynomial time since only one label can be present for each customer at each iteration. However, the strictness of dominance rules is gradually increased as columns become harder to find. This scheme is generally used to find columns quickly and stabilize the dual variables. Once dual variables stabilize stronger heuristics are used, see Desaulniers et al. (2006). When the labeling algorithm fails to find a negative-reduced-cost column with $j = 0$, we increase j by one, i.e., $j = 1$. Now, the labeling algorithm uses dominance rules based on distance and time, set by Algorithm 2. If this fails to find a negative-reduced-cost column, we use dual variable costs based on whether the total load exceeds a the OD-vehicle capacity, as shown in algorithm 2. Finally, if Heuristic pricing fails, we use the exact pricing algorithm with ng-routes relaxations and DSSR.

Algorithm 2: Dominance rules

Input: $j, \mathcal{L}_1, \mathcal{L}_2$

if $j = 0$ **then**
  Set the following dominance rules:
  $d_1 \leq d_2$
else if $j = 1$ **then**
  Set the following dominance rules:
  $t_1 \leq t_2$
  $d_1 \leq d_2$
else
  Set the following dominance rules:
  $t_1 \leq t_2$
  $d_1 \leq d_2$
  if $l_1 \leq Q'$ **then**
    $\bar{\pi}_1 \geq \bar{\pi}_2$
  else
    $d_1 - \bar{\pi}_1 \leq d_2 - \bar{\pi}_2$

4.4 Column Generation Heuristic

Commercial solvers, e.g, CPLEX, are efficient at solving set covering problems. The column generation heuristic consists in adding variables to the restricted master problem until no more variables can be found, and then, solving the resulting MIP with a commercial solver, i.e., without generating more columns.

Using the B&P framework, we can explore nodes in the branch-and-bound tree without providing exact solutions for the pricing problem. Heuristic pricing generates all the columns used and the exact pricing method is not employed. Once no more columns are generated, we branch following the criteria in section 4.2.2. When we have explored $H \in \mathbb{N}$ nodes of the tree, we send the MIP to a commercial solver. The solver provides a best solution given the set of routes in the restricted master problem. Algorithm 3 describes Col-Gen-H.

Algorithm 3: Col-Gen-H

Input: H

$k \leftarrow 0$;
while $k \leq H$ **do**
  $k \leftarrow k + 1$;
  Select node with best bound;
  Heuristic pricing;
  Branch, see section 4.2.2;
Solve restricted master problem MIP.

5 Computational results

In this section, we perform extensive computational experiments to evaluate the solution approach and provide valuable insights into the properties of crowd-shipping. All algorithms were implemented in Java SE 1.8.0 and executed in a Linux-CentOS 7 system with an Intel core E5-2683 at 2.1GHz, and 16GB of ram. The commercial solver CPLEX 12.9 provided by IBM was used to solve the restricted master problem.

In Section 5.1, we explain how we generate the instances that are used for the computational experiments. In Section 5.2, we look at the performance of the proposed heuristics compared with the exact B&P algorithm. In Section 5.3, we examine the changes that occur in the solution when the capacity of ODs is reduced or increased. In Section 5.4, we look at how changes to the route duration constraints affect the results, and finally, in Section 5.5, we perform a sensitivity analysis on the cost structure of the problem and gain valuable insights about the destination of ODs.

5.1 Instance generation

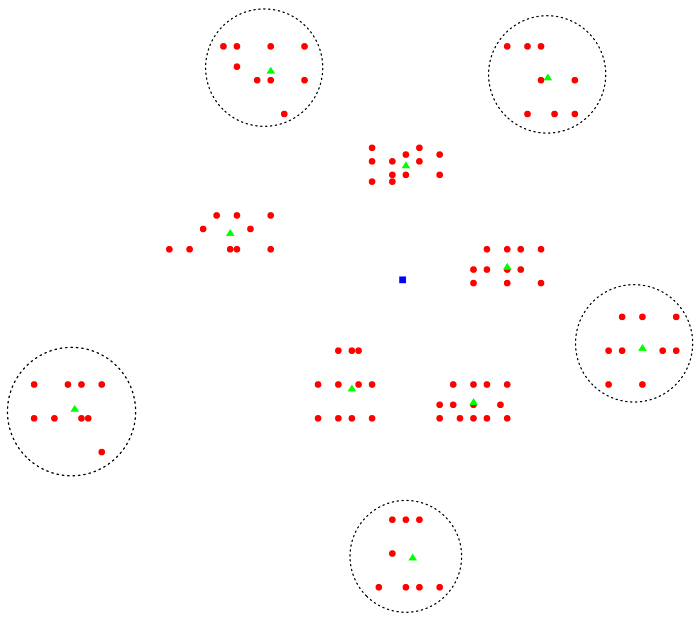
The set of instances is created by modifying the well-known Solomon instances C1 with 25, 50 and 100 customers. The C1 set of instances with 100 customers has 10 clusters that we use to represent the sectors in a city $u \in U$, i.e., $|U| = 10$ when we consider all 100 customers. The discrete probability function used for all instances is the binomial distribution $\xi_u \sim B_u(p_u, M_u)$ for each sector $u \in U$, where p_u is the probability of success of each trial and M_u is the total size of the pool of ODs. The main set of instances are a base case C1-B, in which the probability p_u is the same for all sectors.

Parameters	**Description**	**Value**
Q	Capacity of PV	200
Q'	Capacity of OD	50
F	PV fixed cost	100
F'	OD fixed cost	25
α	Second stage penalty	2.0
β'	OD variable cost	0.25
p_u	Probability of success in binomial distribution	0.04
M_u	Size of the pool of ODs	100
$p_u \times M_u$	Average number of ODs	4
τ	Fixed route duration value	30
ϕ	Variable route duration w.r.t. distance	1.5

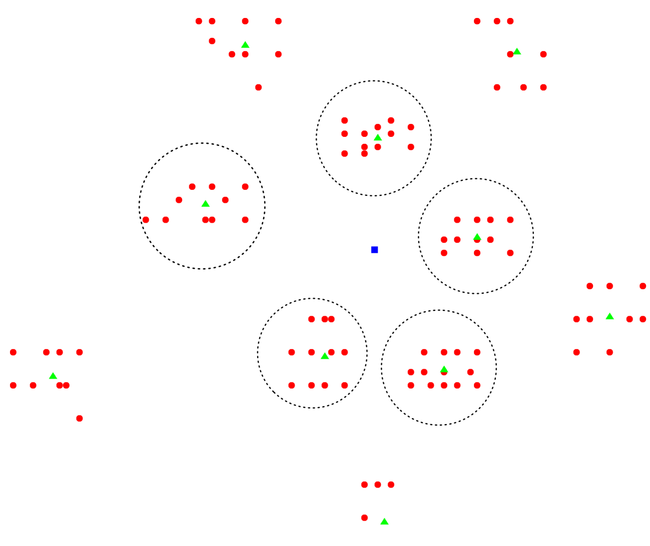
Table 2: Parameters for the base case (C1-B)

Table 2 shows the values for the parameters in the base case C1-B. PVs have the capacity that is provided in the Solomon instances, i.e., $Q = 200$, with an additional fixed cost of 100 and a variable cost that equals the distance. We consider that M_u equals 100 for all sectors and instances, and consider an average value of 4 vehicles for each sector. ODs' vehicles are considerably smaller than

PVs. Hence, to keep the costs consistent and proportional with the capacity, the fixed and variable costs for ODs are one fourth of the costs of PVs, i.e., $F' = 25$, $\beta' = 0.25$, and the capacity of ODs is 50 instead of 200.



(a) Far sectors (C1-F)



(b) Near sectors (C1-N)

Figure 3: Instances C1-F and C1-N

In addition to the base case instance C1-B, described in table 2, we create two other instances C1-F and C1-N. Figure 3a shows the five sectors that are furthest from the depot, while figure 3b shows the five sectors that are nearest to the depot. In instances C1-F, ODs only have destinations in sectors that are far from the depot as shown in 3a, while in instances C1-N ODs only have destinations is sectors that are near the depot, as shown in 3b. Notice that ODs are free to perform visits to customers that are in different sectors regardless of their destination so long as the constraints are satisfied, i.e., capacity, time window, and route duration constraints. However, C1-B instances have higher participation since both near and far destinations are considered. In an effort to adjust the imbalance we increase the participation by increasing the probability of the binomial distribution for instances C1-F and C1-N, i.e., in the probability function, $B_u(p_u, M_u)$, we set $p_u = 0.05$ for Near and Far instances, instead of $p_u = 0.04$ for the Base instance. Since the probability p_u can be interpreted as the participation rate, a higher value indicates that the likelihood of having participation of ODs increases.

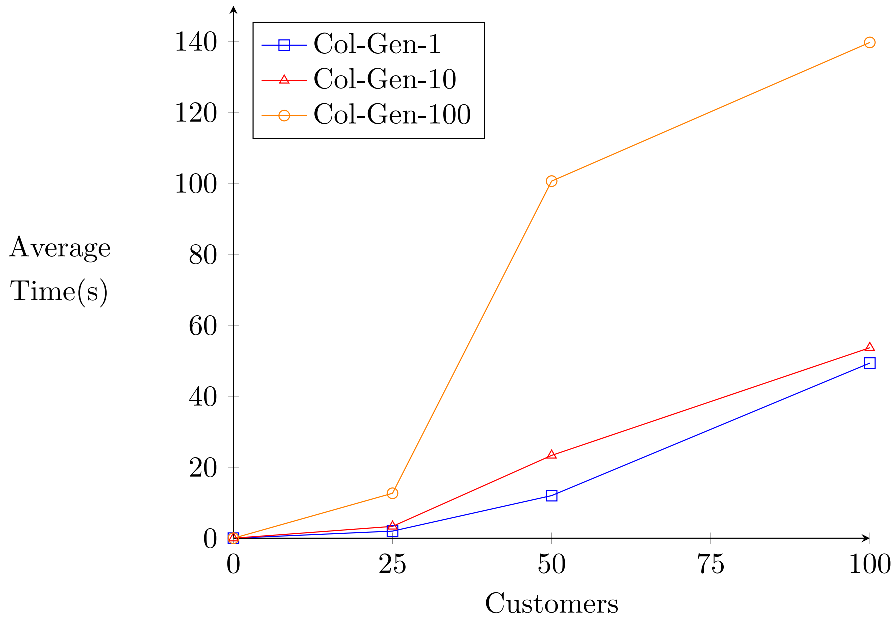
5.2 Performance

Table 3 shows the performance of the exact branch-and-price algorithm and the three heuristics that where implemented for all instances with 25, 50 and 100 customers. In the B&P column, we report the total instances that were solved by B&P, we use bold font when all instances were solved. The two following columns report the lower bound and the time in seconds it took the algorithm to stop. B&P was allowed to run for a maximum of 5 hours; we stop B&P once a feasible solution is found with a gap of less than 1%. B&P was able to solve all instances with 25 customers and most instances with 50 customers, however, no instance with 100 customers was solved. The following columns of table 3 report the results for heuristics C-Gen-1, C-Gen-10 and C-Gen-100. The first column is the upper bound obtained, next is the time in seconds, followed by the gap with respect to the lower bound obtained by B&P.

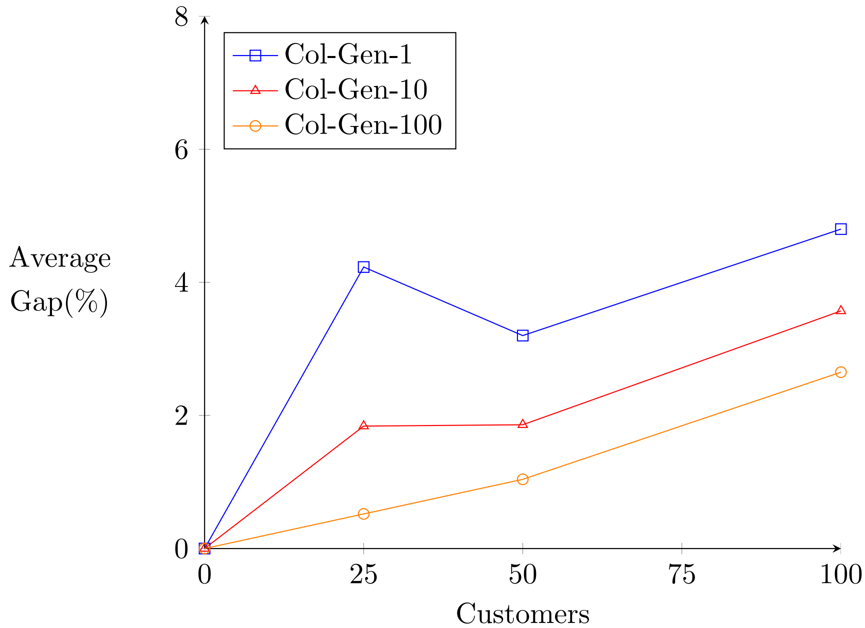
C-Gen-1 is the fastest heuristic: it takes only 2 seconds on small instances and less than a minute on larger instances. However, the average gap is higher than 3% even for small 25-customer instances. We can see in table 3 that the gaps consistently improve as we explore more nodes of the branch-and-bound tree. C-Gen-10 explores 10 nodes of the tree, and thus, we can see major improvements in the gaps while the time it takes to terminate remains competitive with C-Gen-1 and does not increase by much. C-Gen-100 explores 100 nodes of the branch-and-bound tree and has the best results with the lowest gaps: less than 1% for small instances and around 3% for large 100-customer instances. The time for C-Gen-100, however, is visibly higher than C-Gen-10, although still reasonable for a heuristic. The difference in gaps between large instances and small instances is mostly due to the not-optimal lower bound obtained by B&P for 100-customer instances. Since all 25-customers instances were solved by B&P with a gap of at most 1%, it is expected that the lower bounds will be stronger than for instances where B&P fails to terminate, i.e., 100-customer instances. Therefore, we conclude that C-Gen-100 provides good results in reasonable times. For the remaining computational experiments presented in this section, we will only use C-Gen-100 and large instances with 100 customers.

Table 3: Column generation performance

		B&P			**C-Gen-1**			**C-Gen-10**			**C-Gen-100**		
Ins	N	**S**	**LB**	**T(s)**	**UB**	**T(s)**	**G(%)**	**UB**	**T(s)**	**G(%)**	**UB**	**T(s)**	**G(%)**
	25	**9**	428.06	623	442.77	2	3.44	430.33	3	0.53	429.6	13	0.36
C1-B	50	7	787.26	1187	819.52	13	4.1	801.40	29	1.8	796.48	111	1.17
	100	0	1683.19	18000	1770.16	62	5.17	1747.96	54	3.85	1728.4	139	2.69
	25	**9**	432.09	104	450.86	2	4.34	434.99	3	0.68	433.51	9	0.33
C1-F	50	7	795.62	6122	812.78	9	2.16	807.41	18	1.48	802.6	65	0.88
	100	0	1678.96	18000	1778.6	49	5.93	1748.71	53	4.16	1734.92	145	3.33
	25	**9**	418.99	801	446.48	2	4.92	436	4	2.47	429.24	16	0.89
C1-N	50	**9**	773.74	6306	799.49	14	3.33	791.49	23	2.29	782.18	126	1.09
	100	0	1694.61	18000	1749.85	37	3.29	1739.98	54	2.71	1726.88	135	1.94



(a) Time



(b) Gap

Figure 4: Performance

5.3 Capacity

Previously, the capacity for ODs was set at 50, 1/4 of the capacity for PVs, i.e., 1/4 of 200. In this section, we look at different values for the capacity of ODs, increasing and decreasing the capacity to observe changes in the solution. Table 4 shows the results for all 100-customer instances with

C-Gen-100. The rows with bold font, represent the base case where the capacity of ODs is equal to 50. We compare the other rows of table 4 with this base case row. The second column shows the capacity of ODs with values of 25, 50, 75 and 100. Under the column vehicles, we show the average number of ODs, PV, and the total average number of vehicles over all instances. Under costs, the first column shows the total average cost for all instances, next is the standard deviation of the cost, followed by the difference and the percentage with respect to the base case row, i.e., the row where Q' is equal to 50.

Table 4: Capacity

Ins	Q'	**Vehicles**			**Costs**			
		OD	**PV**	**Total**	**Cost**	σ	Δ	%
C1-B	25	0.12	9.88	10.0	1826.04	8.34	97.6	5.65
	50	**9.33**	**7.0**	**16.33**	**1728.4**	**12.36**	**0**	**0**
	75	13.55	4.66	18.21	1591.56	21.1	-136.84	-7.92
	100	19.44	0.0	19.44	1002.65	20.4	-725.75	-41.99
C1-F	25	0.12	9.88	10.0	1826.01	8.46	91.08	5.25
	50	**10.44**	**6.66**	**17.11**	**1734.93**	**14.82**	**0**	**0**
	75	12.33	5.0	17.33	1556.78	14.76	-178.14	-10.27
	100	12.11	3.55	15.66	1201.7	40.4	-533.23	-30.73
C1-N	25	0.12	9.88	10.0	1825.17	10.97	98.28	5.69
	50	**12.11**	**6.22**	**18.33**	**1726.88**	**11.85**	**0**	**0**
	75	12	5.0	17.0	1561.09	15.83	-165.79	-9.60
	100	9.33	5.0	14.33	1392.27	9.54	-334.61	-19.38

When the capacity of ODs increases, we expect to have longer OD-routes that can visit more customers, decreasing the total cost and reducing the number of PVs needed in the solution. Figure 5 shows that the use of ODs in the solution only increases for C1-B instances for the values of Q' from 25 to 100. In fact, the use of ODs decreases for C1-F and C1-N instances when the capacity increases. Figure 6 shows that the total average number of PVs is reduced for instances C1-B and C1-F, but not for C1-N instances.

The fleet of PVs is entirely replaced by ODs for C1-B when $Q' = 100$. This is due to the greater availability of ODs for this instance and the greater number of potential destinations for ODs. In addition, the increase in OD-capacity enables a more significant portion of the demand to be satisfied by ODs. The cost is reduced significantly when the capacity increases.

In instances C1-F, we can see that the total number of ODs starts to decline when the OD-capacity changes from 75 to 100. The total average number of PV continues to decrease in the same range. The cost reduction for C1-F is less than for C1-B, yet larger than the cost reduction for C1-N.

In instances C1-N, we can see that the decrease in ODs starts happening when going from 50 to 75, and the decrease accelerates from 75 to 100, while the PVs are not reduced any further for the range 75 to 100. This finding implies that if the objective is to minimize the use of PVs, then a further increase of OD-capacity is not always guaranteed to reduce the use of PVs.

The cost is reduced for all instances when OD-capacity is increased. The findings in this section show that the use of ODs can be reduced by increasing the OD-capacity, while PV-routes are not necessarily reduced. Under some circumstances, logistic companies might find it useful to reduce the number of PVs used or increase the participation of crowd-shippers to guarantee a larger fleet of ODs. By increasing OD-capacity, we are excluding all crowd-drivers that lack the OD-capacity required, reducing the flexibility and participation in the platform. However, targeting only ODs with large vehicles reduces the cost and reduces the use of the professional fleet in most cases. The platform needs to evaluate if an increase in OD-capacity will indeed result in a reduction of PVs e.g., increasing the OD-capacity from 75 to 100 in instances C1-N does not reduce the use of PVs. The framework provided here can be useful to make such decisions and assess their impact.

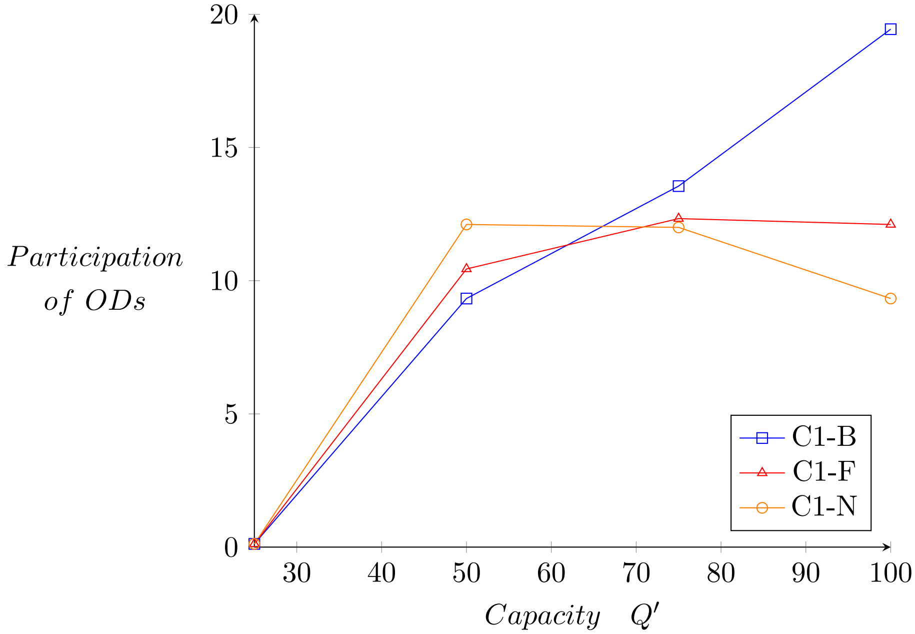


Figure 5: Participation vs capacity of ODs

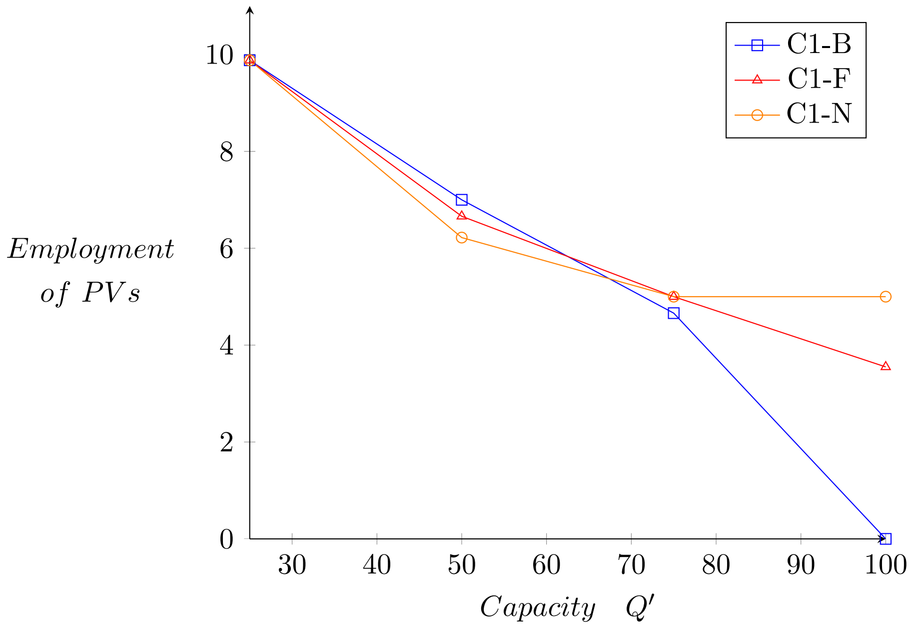


Figure 6: PV employment vs capacity of ODs

5.4 Route duration constraints

Route duration constraints are defined with respect to two different parameters, i.e., ϕ and τ. Table 5 follows the same structure as Table 4 with the addition of the route duration parameters. While ϕ is proportional to the expected trajectory of OD, τ is a value added to the expected trajectory. The difference between the two might seem inconsequential, however, table 5 shows that they have a different impact on instances, particularly C1-F and C1-N. Increasing the value of ϕ leads to cost reductions for all instances; however, the impact on C1-F and C1-B instances is more noticeable than increasing τ. On the other hand, increasing the value of τ leads to a more important reduction of the cost in instances C1-N.

When OD-capacity is equal to 50, the route duration constraint is not a big factor in the cost. The cost increases by at most 5% when the route duration constraint is more restrictive. However, when the capacity is increased to 100 the route duration constraint plays a more important role. The OD-capacity and the route duration constraint are thus connected. If the CSP increases the OD-capacity without increasing the route duration constraint, little savings can be achieved. However, by increasing both to appropriate levels, the cost can be reduced by a significant amount. CSPs can find the best combination of these parameters by utilizing the framework provided in this paper.

Table 5: Route duration constraint

Ins	Parameters			Vehicles			Costs			
	ϕ	τ	Q'	OD	PV	Total	Cost	σ	Δ	%
C1-B	**1.5**	**20**	**50**	**9.33**	**7.0**	**16.33**	**1728.4**	**12.36**	**0**	**0**
	1.0	5	50	9.0	2.77	11.77	1811.28	25.6	82.88	4.80
	1.0	10	50	7.66	7.44	15.11	1770.53	25.48	42.13	2.44
	1.0	20	50	9.44	6.88	16.33	1737.34	17.92	8.94	0.52
	1.25	0	50	3.0	8.88	11.88	1802.43	30.90	70.03	4.28
	1.5	0	50	8.66	7.11	15.77	1745.71	21.95	17.04	1.00
	2.0	0	50	9.0	7.0	16.0	1735.1	13.42	6.69	0.39
	1.5	**20**	**100**	**19.44**	**0.0**	**19.44**	**1002.64**	**20.40**	**0**	**0**
	1.0	5	100	4.33	8.11	12.44	1756.93	86.21	754.28	75.23
	1.0	10	100	10.33	5.33	15.66	1568.53	99.92	565.88	56.44
	1.0	20	100	20.0	0.0	20.0	1099.56	43.88	96.9	9.67
	1.25	0	100	6.66	6.77	13.44	1656.73	114.91	654.07	65.24
	1.5	0	100	15.77	2.33	18.11	1301.33	77.77	298.68	29.79
	2.0	0	100	19.44	0.11	19.55	1054.55	32.12	51.9	5.18
C1-F	**1.5**	**20**	**50**	**10.44**	**6.66**	**17.11**	**1734.93**	**14.82**	**0**	**0**
	1.0	5	50	3.33	9.11	11.55	1813.03	29.12	78.10	4.5
	1.0	10	50	3.44	8.77	12.22	1793.93	33.57	59.0	3.04
	1.0	20	50	7.22	7.55	14.77	1744.96	15.23	10.03	0.58
	1.25	0	50	4.77	8.33	13.11	1794.27	38.33	59.34	3.42
	1.5	0	50	7.22	7.55	14.77	1743.47	15.77	8.54	0.49
	2.0	0	50	7.66	7.44	15.11	1742.1	14.94	7.16	0.41
	1.5	**20**	**100**	**12.11**	**3.55**	**15.66**	**1201.70**	**40.34**	**0**	**0**
	1.0	5	100	3.33	8.44	11.77	1757.86	85.16	556.16	46.28
	1.0	10	100	8.22	6.0	14.22	1601.87	97.50	400.17	33.30
	1.0	20	100	12.11	3.66	15.77	1257.42	63.08	55.62	4.64
	1.25	0	100	8.88	5.66	14.55	1609.45	112.23	407.75	33.93
	1.5	0	100	12.11	3.66	15.77	1253.53	66.27	51.83	4.31
	2.0	0	100	12.55	3.33	15.88	1209.65	44.05	7.95	0.66
C1-N	**1.5**	**20**	**50**	**12.11**	**6.22**	**18.33**	**1726.88**	**11.85**	**0**	**0**
	1.0	5	50	2.55	9.11	11.66	1824.15	12.82	97.26	5.63
	1.0	10	50	3.88	8.55	12.44	1785.47	26.63	58.58	3.39
	1.0	20	50	9.0	7.0	16	1746.03	14.07	19.14	1.11
	1.25	0	50	0.44	9.77	10.22	1823.71	10.13	96.82	5.61
	1.5	0	50	3.0	8.88	11.88	1801.0	28.19	74.12	4.29
	2.0	0	50	7.66	7.44	15.11	1755.92	21.37	29.03	1.68
	1.5	**20**	**100**	**9.33**	**5.0**	**14.33**	**1392.27**	**9.54**	**0**	**0**
	1.0	5	100	1.11	9.44	10.55	1814.2	23.44	421.92	30.30
	1.0	10	100	6.77	7.22	14	1722.09	54.58	329.81	23.69
	1.0	20	100	5.0	9.66	14.66	1399.9	11.45	7.63	0.55
	1.25	0	100	1.44	9.33	10.77	1813.97	23.08	421.69	30.29
	1.5	0	100	7.33	7.0	14.33	1736.17	62.78	343.89	24.70
	2.0	0	100	10.0	5.0	15.0	1441.32	30.43	49.04	3.52

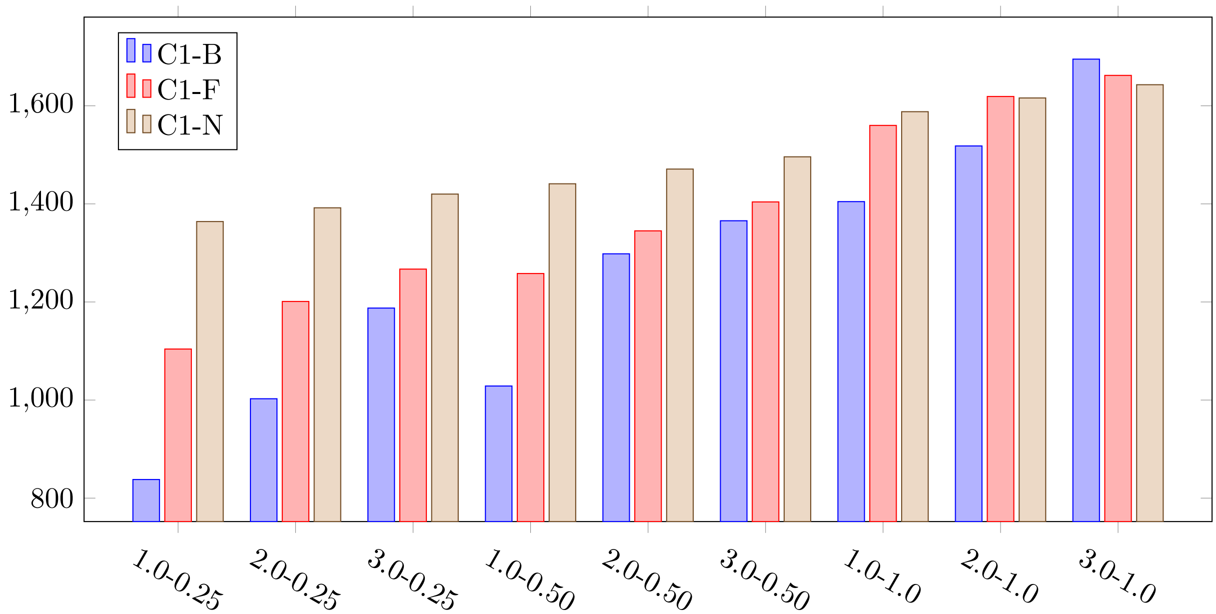


Figure 7: Cost for values of (α-β') for $Q' = 100$

5.5 Cost structure

In this section, we examine the impact of changing the values of parameters α, i.e., the penalty for recourse actions, and β', i.e., the variable cost of ODs. We find that when ODs have destinations that are far from the depot, the cost is reduced the most. However, the cost is more sensitive to an increase of variable cost, i.e., β'. Table 6 follows the same structure as table 4.

The increase in the variable cost makes ODs with destinations near the depot a slightly better option than ODs with destinations far from the depot. In all other cases, ODs with a destination located far from the depot are a better option to reduce the cost. This is more noticeable in instances with an OD-capacity of 100. Longer routes are more expensive with the increase of the variable cost. Interestingly, closer destinations are also less affected when we increase the penalty for the recourse action. This result is due to the return trip that PVs must make to fulfill delivery requests that have been rejected by ODs. Destinations that are far from the depot imply a longer trip back for the PV that is used to mitigate the route rejection. C1-B instances are the most affected by increases in the recourse cost due to the lower value of probability of participation per sector, i.e., $p_u = 0.04$, while in both other instances $p_u = 0.05$ for all sectors. This property can be important in settings where the recourse actions are expensive; in this case, OD-destinations that are near the depot would be more desired. Figure 7 shows the increase of total expected cost of all three instances when the recourse penalty and the variable cost is increased. It shows that C1-N instances are more stable when variable costs are increased, while C1-F instances are the most sensitive to the change in variable cost.

In addition, the proportion of PVs and ODs used in the solution is more stable for C1-N instances. The CSP could keep a stable fleet of PVs at the depot and manage the variable cost of ODs based on demand, without much changes to the fleet.

Table 6: Cost structure

Ins	Parameters			Vehicles			Costs			
	α	β'	Q'	**OD**	**PV**	**Total**	**Cost**	σ	Δ	%
C1-B	1.0	0.25	50	16.55	5.11	21.66	1676.86	11.17	-51.53	-2.98
	2.0	**0.25**	**50**	**9.33**	**7.0**	**16.33**	**1728.4**	**12.36**	**0**	**0**
	3.0	0.25	50	8.87	7.13	16.0	1753.78	17.22	25.38	1.47
	1.0	0.50	50	9.88	6.77	16.66	1778.32	16.70	49.91	2.89
	2.0	0.50	50	5.22	8.11	13.33	1798.12	23.04	69.72	4.03
	3.0	0.50	50	3.44	8.66	12.11	1806.54	22.79	78.14	4.52
	1.0	1.00	50	0.77	9.55	10.33	1819.73	12.50	91.33	5.28
	2.0	1.00	50	0.55	9.66	10.22	1823.47	10.13	95.07	5.5
	3.0	1.00	50	0.33	9.77	10.11	1825.76	8.17	97.36	5.63
	1.0	0.25	100	19.55	0.0	19.55	837.94	18.35	-164.71	-16.43
	2.0	**0.25**	**100**	**19.44**	**0.0**	**19.44**	**1002.64**	**20.40**	**0**	**0**
	3.0	0.25	100	19.33	0.11	19.44	1187.54	20.44	184.89	18.44
	1.0	0.50	100	19.44	0.0	19.44	1028.61	17.9	25.96	2.59
	2.0	0.50	100	19.44	0.0	19.44	1198.19	16.18	195.53	19.50
	3.0	0.50	100	18.77	0.22	19.0	1365.59	22.30	362.94	36.20
	1.0	1.00	100	19.0	0.0	19.0	1404.64	18.28	402	40.09
	2.0	1.00	100	18.77	0.11	18.88	1518.15	21.45	577.49	57.60
	3.0	1.00	100	10.33	4.33	14.66	1695.17	26.96	692.52	69.07
C1-F	1.0	0.25	50	10.77	6.55	17.33	1688.43	12.59	-46.49	-2.68
	2.0	**0.25**	**50**	**10.44**	**6.66**	**17.11**	**1734.93**	**14.82**	**0**	**0**
	3.0	0.25	50	5.44	8.0	13.44	1769.51	18.44	34.59	1.99
	1.0	0.50	50	5.66	8.0	13.66	1778.45	13.58	43.52	2.51
	2.0	0.50	50	3.44	8.66	12.11	1809.49	16.47	74.56	4.30
	3.0	0.50	50	1.55	9.33	10.88	1814.72	17.76	79.79	4.60
	1.0	1.00	50	0.12	9.88	10.0	1827.33	4.52	92.39	5.33
	2.0	1.00	50	0.12	9.88	10.0	1827.58	3.75	92.65	5.34
	3.0	1.00	50	0.12	9.88	10.0	1827.57	3.77	92.64	5.34
	1.0	0.25	100	14.66	2.44	17.11	1104.05	89.25	-97.65	-8.13
	2.0	**0.25**	**100**	**12.11**	**3.55**	**15.66**	**1201.70**	**40.34**	**0**	**0**
	3.0	0.25	100	10.77	4.22	15.0	1267.85	20.56	66.15	5.50
	1.0	0.50	100	13.66	2.77	16.44	1258.83	60.75	57.13	4.75
	2.0	0.50	100	11.44	3.88	15.33	1344.98	27.15	143.29	11.92
	3.0	0.50	100	10.55	4.33	14.88	1404.99	16.19	203.29	16.92
	1.0	1.00	100	11.0	4.12	15.12	1559.95	24.80	358.24	29.81
	2.0	1.00	100	10.44	4.44	14.88	1618.95	15.92	417.25	34.72
	3.0	1.00	100	9.22	5.0	14.22	1662.12	11.40	460.42	38.31
C1-N	1.0	0.25	50	13.0	6.0	19.0	1686.05	16.83	-30.83	-1.79
	2.0	**0.25**	**50**	**12.11**	**6.22**	**18.33**	**1726.88**	**11.85**	**0**	**0**
	3.0	0.25	50	7.55	7.44	15.0	1763.12	18.56	36.23	2.10
	1.0	0.50	50	8.22	7.22	15.44	1771.11	14.10	44.22	2.56
	2.0	0.50	50	5.22	8.11	13.33	1788.9	20.54	62.01	3.59
	3.0	0.50	50	8.33	4.44	12.77	1800.88	21.59	74.0	4.29
	1.0	1.00	50	0.88	9.55	10.43	1817.16	15.77	90.27	5.23
	2.0	1.00	50	0.88	9.55	10.44	1820.51	11.90	93.62	5.42
	3.0	1.00	50	0.88	9.55	10.43	1821.96	9.65	95.07	5.51
	1.0	0.25	100	9.22	5.0	14.22	1364.22	9.04	-28.04	-2.01
	2.0	**0.25**	**100**	**9.33**	**5.0**	**14.33**	**1392.27**	**9.54**	**0**	**0**
	3.0	0.25	100	9.11	5.0	14.11	1420.07	11.40	27.80	2.0
	1.0	0.50	100	9.33	5.0	14.33	1441.29	9.23	49.01	3.52
	2.0	0.50	100	9.33	5.0	14.33	1471.24	12.76	78.96	5.67
	3.0	0.50	100	9.11	5.0	14.11	1495.59	11.22	103.32	7.42
	1.0	1.00	100	9.22	5.0	14.22	1588.13	16.85	195.86	14.07
	2.0	1.00	100	9.22	5.0	14.22	1616.73	19.28	224.46	16.12
	3.0	1.00	100	9.0	5.0	14.0	1643.05	18.02	250.77	18.01

6 Conclusions and future research

In this paper, we extended the framework proposed by Torres et al. (2020) to a setting where OD-destinations are uncertain. We adapted and strengthened the upper bound on the total number of ODs used. We extended the model to this new setting where the specific destination of the supplies of ODs are explicitly considered. Furthermore, a route duration constraint was introduced that promotes participation by keeping routes short. A heuristic was derived from the B&P method that can solve large instances quickly.

Through the numerical experiments conducted, we showed how the use of ODs increases and then starts declining when the OD-capacity is increased. When OD-capacity is high, the total cost decreases, however, the number of ODs needed is reduced. If maintaining a high level of OD-participation is imperative, then a CSP could reduce the required OD-capacity to increase OD-participation. Furthermore, we show how route duration constraints need to be set in relationship with the capacity constraint. If we increase the route duration without increasing the capacity, little cost reductions are obtained. We also found that route duration constraints that consider a proportion of the length of the OD-trajectory are better suited for OD-destinations that are far from the depot. Conversely, route duration constraints that consider a value added to the length of the OD-trajectory have a larger impact on OD-destinations that are near the depot.

In addition, we showed how the geographic location of OD-destinations can affect the total cost. ODs with destinations located further away from the depot have a potential to reduce the cost more than ODs with destinations near the depot. However, this can change in a setting where failed deliveries occur and a return trip to the depot or another action is required (e.g., drones could be used to retrieve undelivered packages). ODs with destinations near the depot will be less inconvenienced by a return trip since their destination is in near proximity to the depot. The literature on stochastic variants largely ignores the possibility of failed deliveries except for Torres et al. (2020). Future research along this line could lead to interesting studies and variants.

Future research that extends this model to other variants is necessary. In general, stochastic variants of crowd-shipping are scarce, albeit, the uncertainty of ODs is an important aspect of the problem. More research that considers the relevant uncertainty is needed, e.g, failed delivery attempts, availability of drivers, or different criteria for accepting routes. In the survey of Alnaggar et al. (2021), the discrepancies between real-world crowd-shipping platforms and the scientific literature are remarked. Specifically, crowd-shipping platforms deliver heterogeneous packages from a central depot with capacitated crowd-drivers. However, few crowd-shipping variants exist in the scientific literature that consider capacitated crowd-drivers and deliveries from a central depot. Studies that consider these important aspects would be good areas of future research.

Acknowledgments

This research was partially funded by the Canadian Natural Sciences and Engineering Research Council (NSERC) under its Discovery Grants research program, as well as the Canada Research Chair in Stochastic Optimization of Transport and Logistics Systems, by Polytechnique Montréal, and by the Ecuadorian Secretaría de Educación Superior, Ciencia, Tecnología e Inovación (SENESCYT) under its international scholarships program.

References

A. Alnaggar, F. Gzara, and J. H. Bookbinder. Crowdsourced delivery: A review of platforms and academic literature. *Omega*, 98:102139, 2021.

AmazonFlex. AmazonFlex, 2021. URL `https://flex.amazon.ca`.

C. Archetti, M. Savelsbergh, and M. G. Speranza. The vehicle routing problem with occasional drivers. *European Journal of Operational Research*, 254(2):472–480, 2016.

C. Archetti, F. Guerriero, and G. Macrina. The online vehicle routing problem with occasional drivers. *Computers & Operations Research*, 127:105144, 2021.

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

B. Asdecker and F. Zirkelbach. What drives the drivers? a qualitative perspective on what motivates the crowd delivery workforce. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020.

R. Baldacci, A. Mingozzi, and R. Roberti. New route relaxation and pricing strategies for the vehicle routing problem. *Operations Research*, 59(5):1269–1283, 2011.

A. Bhatti, H. Akram, H. M. Basit, A. U. Khan, S. M. Raza, and M. B. Naqvi. E-commerce trends during covid-19 pandemic. *International Journal of Future Generation Communication and Networking*, 13(2):1449–1452, 2020.

L. Costa, C. Contardo, and G. Desaulniers. Exact branch-price-and-cut algorithms for vehicle routing. *Transportation Science*, 53(4):946–985, 2019.

L. Dahle, H. Andersson, and M. Christiansen. The vehicle routing problem with dynamic occasional drivers. In *International Conference on Computational Logistics*, pages 49–63. Springer, 2017.

L. Dahle, H. Andersson, M. Christiansen, and M. G. Speranza. The pickup and delivery problem with time windows and occasional drivers. *Computers & Operations Research*, 109:122–133, 2019.

I. Dayarian and M. Savelsbergh. Crowdshipping and same-day delivery: Employing in-store customers to deliver online orders. *Optimization Online*, pages 07–6142, 2017.

G. Debreu. Representation of a preference ordering by a numerical function. *Decision processes*, 3: 159–165, 1954.

G. Desaulniers, J. Desrosiers, and M. M. Solomon. *Column generation*, volume 5. Springer Science & Business Media, 2006.

X. Gao, X. Shi, H. Guo, and Y. Liu. To buy or not buy food online: The impact of the covid-19 epidemic on the adoption of e-commerce in china. *PloS one*, 15(8):e0237900, 2020.

K. Gdowska, A. Viana, and J. P. Pedroso. Stochastic last-mile delivery with crowdshipping. *Transportation research procedia*, 30:90–100, 2018.

M. Gendreau, O. Jabali, and W. Rei. Chapter 8: Stochastic vehicle routing problems. In *Vehicle Routing: Problems, Methods, and Applications, Second Edition*, pages 213–239. SIAM, 2014.

G. Macrina and F. Guerriero. The green vehicle routing problem with occasional drivers. In *New Trends in Emerging Complex Real Life Problems*, pages 357–366. Springer, 2018.

G. Macrina, L. D. P. Pugliese, F. Guerriero, and G. Laporte. Crowd-shipping with time windows and transshipment nodes. *Computers & Operations Research*, 113:104806, 2020.

L. Mrkela and Z. Stanimirović. A variable neighborhood search for the budget-constrained maximal covering location problem with customer preference ordering. *Operational Research*, pages 1–39, 2021.

M. Poggi and E. Uchoa. Chapter 3: New exact algorithms for the capacitated vehicle routing problem. In *Vehicle Routing: Problems, Methods, and Applications, Second Edition*, pages 59–86. SIAM, 2014.

M. Raj, A. Sundararajan, and C. You. Covid-19 and digital resilience: Evidence from uber eats. *arXiv preprint arXiv:2006.07204*, 2020.

G. Righini and M. Salani. New dynamic programming algorithms for the resource constrained elementary shortest path problem. *Networks: An International Journal*, 51(3):155–170, 2008.

A. Sampaio, M. Savelsbergh, L. Veelenturf, and T. Van Woensel. Crowd-based city logistics. In *Sustainable Transportation and Smart Logistics*, pages 381–400. Elsevier, 2019.

D. G. Sanchez, N. G. Parra, C. Ozden, and B. Rijkers. Which jobs are most vulnerable to covid-19? what an analysis of the european union reveals. *World Bank Research and Policy Briefs*, 2020.

F. Torres, M. Gendreau, and W. Rei. Vehicle routing with stochastic supply of crowd vehicles and time windows. *CIRRELT*, 2020.

L. A. Wolsey. *Integer programming.* John Wiley & Sons, 2020.