WORKING PAPER SERIES

Consistent Routing for Local Same-Day Delivery via Micro-Hubs

Charlotte Ackva/Marlin Ulmer

Working Paper No. 7/2022

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

FACULTY OF ECONOMICS AND MANAGEMENT

Impressum (§ 5 TMG)
Herausgeber:
Otto-von-Guericke-Universität Magdeburg
Fakultät für Wirtschaftswissenschaft
Der Dekan

Verantwortlich für diese Ausgabe:
Charlotte Ackva and Marlin Ulmer
Otto-von-Guericke-Universität Magdeburg
Fakultät für Wirtschaftswissenschaft
Postfach 4120
39016 Magdeburg
Germany

http://www.fww.ovgu.de/femm

Bezug über den Herausgeber
ISSN 1615-4274

Consistent Routing for Local Same-Day Delivery via Micro-Hubs

Charlotte Ackva *[1] and Marlin Ulmer[1]

[1]Otto-von-Guericke Universität Magdeburg, Fakultät für Wirtschaftswissenschaft, Management Science

July 5, 2022

Abstract

An increasing number of local shops offer local same-day delivery to compete with the online giants. However, the distribution of parcels from individual shops to customers reduces the rare consolidation opportunities in the last mile even further. Thus, shops start collaborating on urban same-day delivery by using shared vehicles for consolidated transportation of parcels. The shared vehicles conduct consistent daily routes between micro-hubs in the city, serving as transshipment and consolidation centres. This allows stores to bring orders to the next micro-hub, where the parcel is picked up by a vehicle and delivered to the micro-hub closest to its destination – if it is feasible with respect to the vehicle's consistent daily schedule. Creating effective schedules is therefore very important. The difficulty of finding an effective consistent route is amplified by the daily uncertainty in order placements. We model the problem as a two-stage stochastic program. The first stage determines the vehicle schedules. The second stage optimises the flow of realized orders. The goal is to satisfy as many orders per day as possible with the shared vehicles. We propose a multiple scenario approach and suggest problem-specific consensus functions for this framework. We assess the method's performance against an upper bound, a practically-inspired heuristic, and the original consensus function. Our approach clearly outperforms the practically-inspired heuristic and the original consensus function. We observe that collaborative delivery via micro-hubs is worthwhile for delivery time promises of two hours or more. Noticeably, for these service promises, the cost of consistency are surprisingly low.

Keywords — micro-hubs, same-day delivery, routing consistency, two-stage stochastic programming, multiple scenario approach

1 Introduction

Urban areas are facing an increasing amount of parcel transportations. This is reinforced by various factors, first and foremost the expanding e-commerce. In December 2021, Statista published a study estimating the revenue of the e-commerce sector to be about EUR 3.44 trillion worldwide in 2022 (Statista 2020). According to this study, an annual growth of 9.96%, and more than 4.87 million customers purchasing online per year can be expected until 2025. In 2020, about 4.05 billion courier, express, and parcel deliveries were made in Germany according to the German Federal Association of Parcel and Express Logistics (Bundesverband Paket und Expresslogistik e.V. (BIEK) 2021). To participate in the

*Corresponding author. Email: charlotte.ackva@ovgu.de.

e-commerce boom, many local businesses start to offer fast same-day delivery. Goods of local shops can be purchased via an online platform, and shipment is organised by the corresponding shop. Offering such services is challenging as conventional delivery by motorised vans not only exacerbates congestion, emissions and noise, but also causes high costs due to relatively small delivery volumes. Furthermore, the use of large delivery vehicles is not always permitted; especially inner city areas and pedestrian zones show rigid access restrictions. At the same time, customers expect to receive their products more sustainable and faster than ever before. According to the Ecommerce Delivery Benchmark Report 2022, 37.5% of UK shoppers see the speed of delivery as the most important incentive to purchase online (Metapack 2022). As a reaction to this, local shops start collaborating for joint delivery, improving low vehicle fill rates, and driving down transportation costs (Datex 2021). Moreover, "retailers are thinking outside the (big) box to reimagine their 'slow-twitch' supply chains, building for 'fast-twitch' models that can serve customers directly and rapidly" (Burns et al. 2022). New instant delivery models arise, focusing on a small market radius, using satellite stores near areas of high demand (Burns et al. 2022). Such transshipment centres – from now on called micro-hubs – increase consolidation opportunities in the joint delivery process even further.

These micro-hubs are used as dropoff and pickup points for online orders. Shops can bring their goods to the closest micro-hub and customers can pick them up a short time later at a nearby micro-hub. Between the micro-hubs, low-emission vehicles such as cargo-bikes perform consistent tours every day to transport the orders within the network. Examples of similar services can be found in various places, e.g., in several Dutch and Scandinavian cities (Cameron 2022, velove 2022). The vehicles follow a predefined daily schedule visiting micro-hubs at predefined times and in a predefined order. While such consistent tours bring many advantages for shops, customers in drivers, finding effective tours is very challenging given the differences in day-to-day orders. Customers that cannot be served have to be outsourced or served the next day, which is expensive or may lead to dissatisfaction, respectively. Thus, the goal is to find a consistent tour between micro-hubs that maximises the expected amount of delivered parcels per day. In order to provide effective schedules, it is important that they are robust to demand variations in time, respect the pickup and delivery sequence of parcels, and capture the expected daily demand pattern.

The problem can be formulated as a two-stage stochastic program. The first stage of the model aims to find a consistent routing schedule for the delivery vehicle between micro-hubs without the exact demand being known. Given this schedule, the second stage determines the flow of realised parcel orders for a specific day. As solving the problem for realistic sizes is computationally intractable, we propose a multiple scenario approach (MSA) as introduced by Bent and Van Hentenryck (2004). This approach makes use of several possible future demand scenarios and their scenario-dependent solutions. Among those, the MSA chooses that solution showing the most similarity to others. For our problem, evaluating similarity is difficult since individual solutions vary in their pickup and delivery requests, thus showing different routing sequences and arrival times at micro-hubs. Due to this complexity, exact comparison of individual arcs as in Bent and Van Hentenryck (2004) may be too restrictive. To this end, we develop different problem-specific consensus functions designed to detect different structures in the solutions. In a computational study we investigate the effectiveness of the MSA with our different consensus functions. As benchmarks we use an upper bound solution without consistency constraints (i.e. a day-solution where decisions about the routing of the vehicle and the flow of parcels are made simultaneously) and a practically-inspired heuristic solution. For analysing the value of our consensus functions, we compare them to the original one proposed in Bent and Van Hentenryck (2004). In our computational study, we observe that MSA performs better on structured demand and when much time is available to serve parcel requests, than for unstructured demand with tight time limitations. We further show that same day delivery on a full working day with a consistent tour barely looses service quality compared to a daily re-optimised, inconsistent routing policy.

The contributions of this paper are as follows. We are among the first to investigate how micro-hubs can be utilised for collaborative same-day delivery of local shops. We present a stochastic two-stage integer program that finds a consistent tour between micro-hubs for pickup and delivery of parcels in a collaborative local delivery system. We are the first to apply the MSA to a pickup and delivery problem with micro-hubs. In order to uncover structural similarities among the individual solutions we design four new consensus functions making use of problem-specific aspects. The different consensus functions are evaluated in a computational study from which we derive the following insights. We find that MSA with any of the proposed consensus functions outperforms a practically inspired benchmark, while in the vast majority of cases leading to better results as the original consensus function from Bent and Van Hentenryck (2004). Overall, our method performs particularly well for pickup and delivery where tight time limitations are given.

The remaining part of this paper is structured as follows. In Section 2, we present related literature. In Section 3, we define the model for consistent pickup and delivery routing. In Section 4, we introduce the multiple scenario approach. In Section 5, we present setup and results of the computational experiments. We finish our paper with a conclusion and outlook in Section 6.

2 Literature Review

Several aspects have to be considered when planning consistent schedules in local delivery services: pickup and delivery needs to be organised through a two-echelon transportation system with transshipment facilities, while consistency in vehicle's routes is to be maintained. This problem of picking up and delivering parcels during one route is denoted as the vehicle routing problem (VRP) with simultaneous pickup and delivery; a review on such problems can be found in Koç et al. (2020). Introducing transshipment facilities to a pickup and delivery problem leads to a two-echelon logistic system which is usually operated by two fleets of vehicles. In general, this is denoted as the two-echelon vehicle routing problem (2E-VRP). A literature review on such problems is presented by Cuda et al. (2015), and more recently by Jiang and Li (2021) and Sluijk et al. (2022). However, most papers focus on delivery only, assume demand to be deterministic, and consequently route first and second fleet at the same time. In contrast, our model aims to determine a routing schedule for the first fleet only in order to provide consistent pickup and delivery service despite daily varying demand. We therefore refer to different concepts of consistency in Section 2.1. After that, we provide some background on scenario-based solution approaches in Section 2.2.

2.1 Consistent Vehicle Routing

We hence seek to find consistent tours for the fleet serving micro-hubs. Kovacs et al. (2014) provide a survey on consistency in VRPs. They distinguish arrival time, person-oriented and delivery quantity consistency, and provide modelling concepts and solution methodology for each type. In our context, we require the even stronger concept of routing consistency: vehicles should always conduct the very same tour, i.e. visit the same micro-hubs at the same times each day. This captures the pickup and delivery dimension of our problem and allows storekeepers to organise transportation of their orders to corresponding micro-hubs in time for further shipment. Yet, in literature arrival time consistency is the concept closest to this and common modelling approaches are imposing hard or soft constraints, previously assigning time windows to customers or determining routes a-priori. Exemplary publications and different approaches can be found in Kovacs et al. (2014), and more recently in Song et al. (2020). Most relevant to our work are consistent VRPs including pickup and delivery. Zhen et al. (2020) propose a consistent VRP for simultaneous distribution and collection in reverse logistics. Emadikhiav et al. (2020) address the simultaneous pickup and delivery of orders of an instrument-calibration company. The goal is to minimise transportation costs while limiting late deliveries and enforcing consistent arrival times. However, in both papers orders are deterministic.

To deal with stochastic customers, different methods are needed. Two prominent advances are assigning time windows to customers previously or determining tours a-priori which are possibly adapted later to the realised demand. The latter can be applied for stochastic customers as well as stochastic demand while the former is applicable only if customer locations are known in advance and only demand volumes vary from day to day. Assigning time windows is often modelled by a two-stage stochastic programming formulation. Spliet and Gabor (2015) for example assign time windows to each customer at the first stage. They formulate a MIP for this stage with the objective to minimise expected travel costs. Once demand volumes are revealed, a VRP has to be solved meeting the previously determined time windows. Dalmeijer and Spliet (2018) can improve the computational performance of this problem by strengthening the problem formulation by valid inequalities. A discrete variant of the above problem is presented by Spliet and Desaulniers (2015). They further propose an exact branch-price-and-cut algorithm. Spliet et al. (2018) extend the problem to time-dependent travel times and develop a branch-price-and-cut algorithm to solve the problem to optimality. A similar problem of previously assigning time windows to customers on first, and routing vehicles on second stage is proposed by Subramanyam et al. (2018). They additionally consider stochastic travel times and propose a scenario decomposition algorithm to solve the problem.

Precedently assigning time windows is not enough for our pickup and delivery routing problem as we look for a schedule with exact time synchronisation. Since we are facing stochastic demand, we need to determine routes before demand becomes known, e.g. based on stochastic information. This concept of time consistency is called a priori routing or finding master tours. When demand is revealed, these routes are commonly updated using recourse actions, such as skipping customers or restocking at the depot for example. Some common recourse strategies are explained in Kovacs et al. (2014). Often, such problems are modelled as a two-stage stochastic program: at the first stage, an a-priori routing is determined under uncertain demand. At the second stage, uncertainty is revealed and corresponding recourse actions are selected. Reviews on a-priori routing problems and corresponding solution methods can be found in Bertsimas et al. (1990), Campbell and Thomas and Kovacs et al. (2014). We concentrate on the most relevant work for our context in the following. Hvattum et al. (2006) use a multistage stochastic programming formulation to model a VRP with both deterministic and stochastic customers. Recourse strategies are applied repetitively based on a sample scenarios heuristic approach. Sungur et al. (2010) describe a courier delivery problem with stochastic customers and uncertain service times. They offer a multi-objective two-stage program for a variant of the ConVRP with time windows to develop a-priori master tours using a scenario-based solution approach. Uncertain travel times and stochastic customers are also considered in Sampaio et al. (2019), who propose a VRP with roaming delivery locations which they solve with a scenario-based sample average approximation. Angelelli et al. (2017) solve a probabilistic team orienteering problem through a two-stage stochastic program that maximises the expected profit of visited customers. They solve their problem applying a branch-and-cut approach as well as different heuristic methods. There is some recent work on two-stage stochastic programs for VRPs with stochastic demands and recourse actions. Lagos et al. (2019) suggest such a model minimising the expected travel costs. The models proposed by Bernardo and Pannek (2018), Salavati-Khoshghalb et al. (2019) and Florio et al. (2022) additionally aim to minimise the expected costs of recourse actions. Similar to our problem, Crainic et al. (2016) suggest a two-stage stochastic programming formulation for the 2E-VRP with stochastic demands. At the first stage, an urban-vehicle service network design model routes the first fleet and determines the general load of micro-hubs, using an approximation of the routing cost from micro-hubs to customers. The second stage concerns the routing of second fleet vehicles and possible recourse actions for the first fleet. The authors evaluate different recourse strategies through repetitively applying the adjusted plan for each planning period. Their work differs from ours as determining loads of hubs is not part of our problem, further we do not apply recourse strategies since we seek a consistent routing between micro-hubs. Consistent master routes are also determined in the work of Visser and Savelsbergh (2019). They investigate a strategic time slot management problem where master tours and time windows at customer locations are determined simultaneously on first stage, facing uncertain demand. In difference to our problem, assigning time slots instead of precise arrival times is sufficient. Further, recourse actions may be applied after demand realisation in the sense that customers may be skipped if they cannot be served within their time window. A different approach for consistent vehicle tours is used by Orenstein and Raviv (2022). The authors propose an urban parcel pickup and delivery system including so called service points that can serve as consolidation, transshipment, and pickup point for customers or drivers. Customers may be served from several service points, what further increases flexibility in the delivery process. They develop a myopic policy to route stochastically arriving parcels based on given vehicle routes. Vehicle routes are determined a-priori using a math heuristic.

We summarise related literature on two-stage stochastic routing problems in Table 1. For each paper, we classify the type of route consistency, the source of uncertainty, the decisions taken on first and second stage, the objective, as well as the solution method. By master tours (with recourse) we refer to those

Paper	**Consistency**		**Stochasticity**			**1st Stage**		**2nd Stage**		**Objective**		**Solution Method**	
	M	TW	pickup/delivery	demand	time	M	TW	parcel flow	routing	service	costs	MSA	exact
Hvattum et al. (2006)	(✓)		(✓)	✓		(✓)			(✓)				
Sungur et al. (2010)	(✓)		(✓)		✓	(✓)			(✓)	✓			
Spliet and Gabor (2015)		✓		✓			✓		✓		✓		✓
Spliet and Desaulniers (2015)		✓		✓			✓		✓		✓		✓
Crainic et al. (2016)	(✓)			✓		(✓)			✓		✓		
Angelelli et al. (2017)	(✓)		(✓)			(✓)			(✓)	✓			✓
Bernardo and Pannek (2018)	(✓)			✓		(✓)			(✓)		✓		
Dalmeijer and Spliet (2018)		✓		✓			✓		✓		✓		✓
Spliet et al. (2018)		✓		✓			✓		✓		✓		✓
Subramanyam et al. (2018)		✓		✓	✓		✓		✓		✓		✓
Lagos et al. (2019)	(✓)			✓		(✓)			(✓)		✓		✓
Salavati-Khoshghalb et al. (2019)	(✓)			✓		(✓)			(✓)		✓		✓
Sampaio et al. (2019)	(✓)				✓	(✓)			(✓)		✓		
Visser and Savelsbergh (2019)	(✓)	✓		✓		(✓)	✓		(✓)				
Song et al. (2020)			(✓)						✓		✓	✓	
Florio et al. (2022)	(✓)			✓		(✓)			(✓)		✓		✓
our paper	✓		✓			✓		✓		✓		✓	

Table 1: Related literature on two-stage stochastic routing problems. Following abbreviations are used: M = master tour (with recourse), TW = time window assignment.

papers that address routing consistency via finding master tours on first stage. However, recourse actions on second stage lead to changes in these master tours. This is why we place brackets around the check mark in columns "M" (Consistency / 1st stage) if recourse actions are applied. Brackets in the "routing" column of the 2nd stage indicate that re-routing decisions are taken for recourse. No brackets here mean solving an entire routing problem. Finally, some papers consider stochastic customers for pickup and / or delivery. If only one of the two is addressed, this is also indicated by brackets in column "pickup/delivery".

In Table 1, we see that all revised papers finding master tours on first stage later apply recourse strategies. This is fundamentally different from our problem where routing schedules between micro-hubs need to be fixed for any demand realisation. Vehicles between micro-hubs stick to their routes regardless of the daily varying customer orders. Different to any other publication, we determine the flow of realised parcel orders on second stage. Further, none of the presented papers on two-stage stochastic routing problems includes both pickup and delivery customers. We hence propose a novel two-stage stochastic program for consistent pickup and delivery at micro-hubs with stochastic customer demand, but without recourse actions. To the best of our knowledge, this problem has not been studied in literature yet.

2.2 Scenario-Based Solution Approaches

To solve this problem, we use a scenario-based solution approach since this is inherent to two-stage stochastic programming. We apply a multiple scenario approach (MSA) as introduced by Bent and Van Hentenryck (2004). The MSA provides a general framework to solve dynamic stochastic multi-stage programs by solving a sampled set of representative deterministic problems. Among those individual solutions, MSA chooses the best one according to a consensus function in order to find a solution that can be expected to perform well for different demand realisations. The original consensus function proposed by Bent and Van Hentenryck (2004) compares the customers that vehicles are scheduled to visit next. Ghiani et al. (2012) use a similar consensus function in their sample scenario planning approach. They solve an uncapacitated travelling salesman problem and compare the position of customers in different routes as consensus function. Song et al. (2020) introduce a new consensus function for a two-stage stochastic assignment and team-orienteering problem that measures the differences between two assignment variables. Some further publications using the MSA are listed in Ritzinger et al. (2016). Most authors applying MSA use consensus functions particularly adapted to their problem structure. That way, they often compare assignment or routing details, for example. Our problem involves both pickup *and* delivery, and hence shows to be much more complex. For this reason, we design further consensus functions in Section 4, comparing different space and time attributes of the individual solutions. In Section 5, we show in which cases they are most effective.

There are several works considering further scenario-based solution approaches. Sungur et al. (2010) for example determine an a-priori routing with a scenario-based approach using historical data. They base their master plan on representative customers most likely to appear, and worst case service times. That is, the information available from the various scenarios is reduced. Instead, using the MSA we consider full information from all scenarios. Azi et al. (2012) use multiple scenario generation for a dynamic VRP to decide about the acceptance of newly arising customer requests. A new request is evaluated among all generated scenarios. If on average, serving this customer leads to a benefit, she is accepted. Another sampling-based solution method is used by Bernardo and Pannek (2018). Different from MSA, scenarios are only sampled at beginning of the planning stage. Then, a static scenario is defined as a weighted linear combination of all sampled scenarios. This instance is then solved using different heuristic approaches. Further scenario-based solution methods can be found in Subramanyam et al. (2018), Sampaio et al. (2019) or Visser and Savelsbergh (2019), for example. A review on different solution methods for stochastic dynamic VRPs, incuding scenario-based approaches, is provided by Soeffker et al. (2021).

3 Model

In this section, we give a detailed description of the consistent pickup and delivery problem with micro-hubs and stochastic customer demand. To that end, we first state the problem in Section 3.1 and highlight the two stages with an illustrative example in Section 3.2. In Section 3.3, we provide the mathematical framework for the model.

3.1 Problem Statement

We propose a model to determine a consistent routing schedule for shared vehicles in collaborative urban delivery. The use of shared vehicles increases consolidation opportunities since parcels from different stores can be bundled for joint delivery to the same region.

To pickup and deliver parcels, a shared vehicle conducts service between micro-hubs that are placed at fixed locations in the city. micro-hubs may either be located in selected stores, or close to customer's locations or shopping areas. The schedule at micro-hubs must be known previously so that stores can organise transport to the micro-hub accordingly. In our system, we use a fixed number of vehicles with a given, finite capacity each. While operating, vehicles are allowed to wait at micro-hubs in order to include later parcels. Also, vehicles are allowed to perform pickup and delivery on the same route and simultaneously during one stop. Moreover, we consider a limited planning horizon within which service is operated. We call this the service time horizon. At the beginning of the service time horizon, all vehicles are located at a depot at the outskirts of the city. We consider a transfer time at each micro-hub a vehicle visits on its itinerary to load and unload parcels. For simplicity, the transfer times are incorporated in the travel times.

We assume that each parcel order consists of a pickup location (store) with a release time, a delivery location (customer), and a homogeneous parcel volume. Further, we assume that each storekeeper brings a parcel to her closest micro-hub as soon as the order is placed. Similarly, when a parcel reaches its final micro-hub, the customer picks it up there directly. To represent this in our model, pickup and delivery locations are mapped to their closest micro-hubs such that each parcel has a pickup micro-hub and a delivery micro-hub. The release time of a parcel indicates the earliest time it can be picked up. We set a fixed delivery time promise to all orders, restricting the difference between release time and delivery time at the parcel's destination micro-hub. We do not have to serve all orders placed, but aim to pickup and deliver as many parcels as possible. Each parcel that is picked up must also be delivered in time.

We model this problem as a two-stage program. The first stage develops the consistent routing schedule for the vehicles, i.e. a sequence of stops at micro-hubs with corresponding departure times. Once this routing is fixed, realised parcel orders are routed according to the given schedule at the second stage. At this, parcels cannot move independently in the network, but must be transported by a vehicle. The goal of the model is to find a routing schedule that maximises the expected amount of daily delivered parcels.

3.2 Example

We illustrate our problem setting in a short example. Assume there are one shared vehicle, a depot and two micro-hubs in the city, as displayed in Figure 1. We further assume that the vehicle has the following first-stage schedule. It leaves the depot at 12:30 and reaches micro-hub 1 at 13:30. After a transfer time of ten minutes it continues to micro-hub 2 where it arrives at 13:50. Again, ten minutes are needed for transfer such that the vehicle leaves micro-hub 2 at 14:00 and reaches the depot at 14:35. This routing schedule is shown in Figure 1. Now, suppose in scenario 1 there are two parcels which both have to be transported from micro-hub 1 to micro-hub 2, indicated by the minus for pickup, and the plus for delivery. The release times are 11:00 for parcel 1 and 13:00 for parcel 2. With the given schedule, both parcels can be picked up and delivered in time.

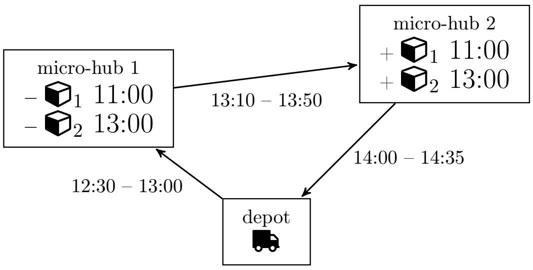


Figure 1: Scenario 1: all parcels can be transported.

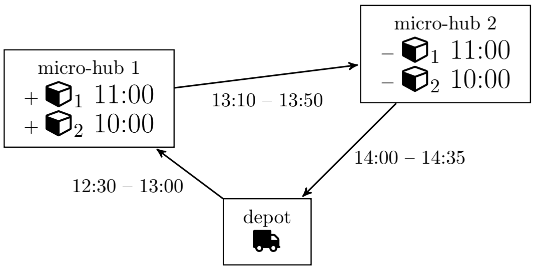


Figure 2: Scenario 2: no parcel can be transported.

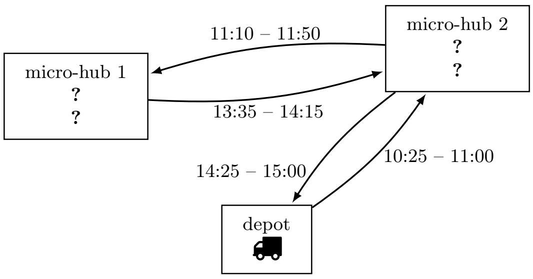


Figure 3: Combined routing schedule.

However, in a different demand scenario, this schedule might be inefficient. For instance, suppose that in a second scenario 2 again two orders are placed, differing in release time, pickup and delivery location: parcels 1 and 2 have to be picked up at micro-hub 2, and brought to micro-hub 1, having release times 11:00 and 10:35, respectively. Note that in this scenario no order can be fulfilled using the given routing schedule. This small example shows that the performance of our proposed pickup and delivery service highly depends on the shared vehicle's routing schedule. To transport as many orders as possible, it thus is very important to create effective schedules that are flexible with respect to order uncertainty in time and space. For the example at hand, a routing schedule allowing to meet the demand in both scenario is presented in Figure 3. The vehicle visits micro-hub 2 first, then continues to micro-hub 1. Instead of returning to the depot directly, it waits there for further parcels to be picked up and then goes back to micro-hub 2. This way, parcels can be transported from micro-hub 1 to micro-hub 2 and vice versa.

3.3 Mathematical Formulation

In this section, we introduce the required notation in Section 3.3.1, and present the stochastic two-stage model in Section 3.3.2. The corresponding deterministic second-stage model is stated in Section 3.3.3. We note that while in our computational study, we focus on the single-vehicle case, in this section, we present the more general model allowing for a fleet of vehicles.

3.3.1 Notation

In the following, we introduce the notation of our model, also summarised in Table 2. Parcel orders are placed on a daily basis. We hence consider a set of several daily scenarios denoted by S. We assign a certain probability of occurrence p_s to each scenario $s \in S$. The set of parcel orders placed on scenario $s \in S$ is denoted by P_s. We refer to an element $p \in P_s$ as parcel or order equivalently. For our model, we consider a set of physical nodes $V = V_0 \cup V_H$, where $V_0 = \{v_0\}$ denotes the depot and V_H represents the set of micro-hubs. The pickup hub (origin) of a parcel $p \in P_s$ is represented by node $o_p \in V_H$, the delivery hub (destination) by node $d_p \in V_H$. The parcel's release time is denoted by r_p for every $p \in P_s$ and indicates the start of a time window of length T_P, within which a parcel $p \in P_s$ must be picked up at micro-hub o_p and delivered to micro-hub d_p in order to fulfil the order. Otherwise, we say that the parcel is not served or that the order is not fulfilled by our system, but must be outsourced. For simplicity we assume that each parcel has a homogeneous volume of 1. Service is operated by a homogeneous fleet of vehicles K. Each vehicle has the same velocity v_K and maximum capacity C_K. Also, micro-hubs have a limited storage capacity C_S. In this section, we present the general model for several vehicles. In our computational study in Section 5, we focus on one vehicle only.

We model our problem over a discrete time horizon $T = \{0, 1, ..., T_{max}\}$ representing one working day. Service time of the vehicle starts at time step 0 and ends at T_{max}. Inspired by the work of Neumann-Saavedra et al. (2016), we make use of a time extended network for the formulation of our problem. In such, nodes are duplicated per time step and arcs are constructed correspondingly. To this end, let $t_{i,j}$ denote the integer travel time between two nodes $i,\ j \in V_0 \cup V_H,\ i \neq j$. Then, we extend the set of physical nodes in the following way: For each node $i \in V_0 \cup V_H$ we create one duplicate per time step $t \in T$ and denote this node as (i,t), $\forall i \in V_0 \cup V_H$, $\forall t \in T$. With these duplicated nodes, we can define the set of arcs on which any vehicle $k \in K$ is allowed to travel as:

$$A_K := \left\{ \big((i,t),(j,\bar{t})\big) \ \middle|\ \bar{t} = t + t_{i,j},\ \forall i,j \in V_0 \cup V_H,\ \forall t, \bar{t} \in T:\ \bar{t} > t \right\}.$$

Furthermore, we define the set of arcs on which parcels $p \in P_s$ for all $s \in S$ are allowed to travel as:

$$A_P := \left\{ \big((i,t),(j,\bar{t})\big) \ \middle|\ \bar{t} = t + t_{i,j},\ \forall i,j \in V_H,\ \forall t, \bar{t} \in T:\ \bar{t} > t \right\}.$$

That is, parcels can be transported between any micro-hubs but are not allowed to enter the depot. Further note that parcels can only travel along arcs when loaded onto a vehicle. Finally, we define the set of waiting arcs as those arcs linking the same micro-hub between two following time steps:

$$A_W := \{((i,t),(i,t+1)) \ \mid \forall i \in V_H,\ \forall t, t+1 \in T\},$$

and the waiting arcs at the depot are defined by:

$$A_{V_0} := \{((0,t),(0,t+1)) \ \mid \forall t, t+1 \in T\}.$$

Note that arcs only go forward in time by construction.

Notation	Description
S	set of scenarios
p_s	probability of occurence of scenario $s \in S$
$T = \{0, 1, .., T_{max}\}$	time horizon in equidistant time steps
$V_0 = \{v_0\}$	depot
V_H	set of micro-hubs
P_s	set of parcels in scenario $s \in S$
K	set of vehicles
C_K	maximum vehicle capacity
C_H	maximum micro-hub capacity
T_P	length of the time window for orders
r_p	release time of parcel $p \in P_s$
o_p	pickup hub of parcel $p \in P_s$
d_p	delivery hub of parcel $p \in P_s$
v_K	vehicle velocity
$t_{i,j}$	travel time between nodes i and j as an integer value, $i, j \in V_0 \cup V_H$
A_K	set of arcs allowed for vehicles
A_P	set of arcs allowed for parcels
A_W	set of waiting arcs between micro-hubs
A_{V_0}	set of waiting arcs at depot

Table 2: Parameters and sets for the model.

Decisions are made on two stages. The first stage concerns the long-term planning of vehicle routes. Although demand varies from day to day, vehicles are to follow a fixed, consistent schedule. For this, we introduce the first stage decision variables $x^k_{(i,t),(j,u)}$ for each $k \in K$. It is defined as:

$$x^k_{(i,t),(j,u)} := \begin{cases} 1, & \text{if vehicle } k \text{ travels directly from node } (i,t) \text{ to node } (j,u), \\ 0, & \text{otherwise.} \end{cases}$$

The second stage concerns the operational daily planning of which parcels to serve. Therefore, we use a binary second stage decision variable $a_{p,s}$ for all $p \in P_s$ to decide whether parcel p is served in scenario $s \in S$ or not:

$$a_{p,s} := \begin{cases} 1, & \text{parcel } p \text{ is served in scenario } s \in S, \\ 0, & \text{otherwise.} \end{cases}$$

We further consider binary second stage decision variables for parcels to decide which arcs they use on their itineraries:

$$z^p_{(i,t),(j,u),s} := \begin{cases} 1, & \text{if parcel } p \in P_s \text{ travels directly from node } (i,t) \text{ to node } (j,u) \text{ in scenario } s \in S, \\ 0, & \text{otherwise.} \end{cases}$$

The objective of the model is to maximise the expected number of delivered parcels.

3.3.2 The Stochastic Two-Stage Model

With the notation above we now state the two-stage stochastic integer program for consistent routing in collaborative urban delivery as follows. Let x, z_s, and a_s be the vectors with entries defined as above.

Then we are looking for a solution of the form $\big(x, (z_s, a_s)_{s\in S}\big)$ to the two-stage stochastic program:

$$\max \sum_{s\in S} p_s \sum_{p\in P_s} a_{p,s} \tag{2-$\mathcal{SP}$}$$
$$\text{s.t.} \quad (x, z_s, a_s) \in C_s \quad \forall s \in S,$$

where C_s represents the feasible set corresponding to scenario s defined by Constraints (1) to (19) following below. The objective of (*2-$\mathcal{SP}$*) maximises the expected amount of parcels that can be served, according to the probability of occurence of each scenario. Note that the first-stage decision variable x is invariant with regard to the resulting scenario. Thus, the following constraints ensure a feasible routing for the vehicles for any demand realisation. Constraints (1) and (2) ensure that each vehicle starts and ends its tour at the depot. That a vehicle can only start its tour there, and not at any micro-hub, is guaranteed by Constraints (3). Through Constraints (4) each vehicle leaves the depot at most once, i.e. vehicles do not return to the tour during their route. Together with the flow conservation constraints in Constraints (5), this prohibits vehicles to visit the depot during service.

$$\sum_{((0,0),(j,u))\in A_K} x^k_{(0,0),(j,u)} = 1 \qquad \forall k \in K, \tag{1}$$

$$\sum_{((i,t),(0,T_{max}))\in A_K} x^k_{(i,t),(0,T_{max})} = 1 \qquad \forall k \in K, \tag{2}$$

$$\sum_{((i,0),(j,u))\in A_K} x^k_{(i,0),(j,u)} = 0, \qquad \forall i \in V_H, \forall k \in K, \tag{3}$$

$$\sum_{((0,t),(j,u))\in A_K\setminus(A_W\cup A_{V_0})} x^k_{(0,t),(j,u)} \leq 1 \qquad \forall k \in K, \tag{4}$$

$$\sum_{((i,t),(j,u))\in A_K} x^k_{(i,t),(j,u)} = \sum_{((j,u),(i,t))\in A_K} x^k_{(j,u),(i,t)} \quad \forall j \in V_H \cup V_0, \forall k \in K, \forall u \in T\setminus\{0, T_{max}\}. \tag{5}$$

The following Constraints (6) to (10) are concerned with the routing of parcels, that does depend on the resulting scenario. For this reason, all following constraints must be kept for any demand realisation $s \in S$. Constraints (6) ensure that each parcel has to start its itinerary at its pickup micro-hub. In order to avoid unnecessary shipment, Constraints (7) guarantee that each micro-hub is visited at most once by a parcel. In the case where pickup and delivery customer are mapped to the same micro-hub, no transportation by vehicle is needed. This is captured by Constraints (8). Constraints (9) and (10) state that each parcel that is served must leave its pickup hub and enter its delivery hub. Note that with this formulation parcels must leave their pickup hub at time $t = 0$ and enter their delivery hub at time $t = T_{max}$. However, this may be satisfied via waiting arcs such that physical leaving and entering may happen later and earlier, respectively. Constraints (11) state the flow conservation constraints of parcels at micro hubs.

$$\sum_{((i,0),(j,u))\in A_P, i\neq o_p} z^p_{(i,0),(j,u),s} = 0 \qquad \forall p \in P, \forall s \in S, \tag{6}$$

$$\sum_{((i,t),(j,u))\in A_P\setminus A_W} z^p_{(i,t),(j,u),s} \leq 1 \qquad \forall j \in V_H, \forall p \in P, \forall s \in S, \tag{7}$$

$$z^p_{(i,t),(j,u),s} = 0 \qquad \forall (i,t),(j,u) \in A_P\setminus A_W, \forall p \in P : o_p = d_p, \forall s \in S, \tag{8}$$

$$\sum_{\big((o_p,0),(j,u)\big)\in A_P} z^p_{(o_p,0),(j,u),s} = a_{p,s} \qquad \forall p \in P, \forall s \in S, \tag{9}$$

$$\sum_{\big((i,t),(d_p,T_{max})\big)\in A_P} z^p_{(i,t),(d_p,T_{max}),s} = a_{p,s} \qquad \forall p \in P, \forall s \in S, \tag{10}$$

$$\sum_{((i,t),(j,u))\in A_P} z^p_{(i,t),(j,u),s} = \sum_{((j,u),(i,t))\in A_P} z^p_{(j,u),(i,t),s} \quad \forall j \in V_H, \forall p \in P, \quad \forall u \in T\backslash\{0, T_{max}\}, \forall s \in S. \tag{11}$$

Parcels can only travel along some arc if some vehicle travels along it, as they cannot move independently but have to be transported by vehicles. To this end, Constraints (12) link the routes of parcels to those of vehicles.

$$z^p_{(i,t),(j,u),s} \leq \sum_{k\in K} x^k_{(i,t),(j,u)} \quad \forall\,((i,t),(j,u)) \in A_P\backslash A_W, \quad \forall p \in P, \forall s \in S. \tag{12}$$

To not exceed vehicle capacity constraints, we have to control the maximum load capacity on vehicle arcs, which is done with Constraints (13). Limited micro-hub capacity is controlled via restricting the load on waiting arcs through Constraints (14).

$$\sum_{p\in P} z^p_{(i,t),(j,u),s} \leq C_K \cdot \sum_{k\in K} x^k_{(i,t),(j,u)} \quad \forall\,((i,t),(j,u)) \in A_P\backslash A_W, \quad \forall s \in S, \tag{13}$$

$$\sum_{p\in P} z^p_{(i,t),(j,u),s} \leq C_H \quad \forall\,((i,t),(j,u)) \in A_W, \quad \forall t \in T\backslash\{0, T_{max}\}, \forall s \in S. \tag{14}$$

Constraints (15) and (16) ensure that no parcel is transported between micro-hubs outside its delivery time window, i.e. not before the parcel is placed nor later than delivery time promise.

$$z^p_{(i,t),(j,u),s} = 0 \quad \forall\,((i,t),(j,u)) \in A_P\backslash A_W, \quad \forall t \in T : t \leq r_p, \forall p \in P, \quad \forall s \in S, \tag{15}$$

$$z^p_{(i,t),(j,u),s} = 0 \quad \forall\,((i,t),(j,u)) \in A_P\backslash A_W, \quad \forall u \in T : u \geq r_p + T_P, \quad \forall p \in P, \forall s \in S, \tag{16}$$

Finally, Constraints (17) to (19) state the domain of the decision variables.

$$x^k_{(i,t),(j,u)} \in \{0,1\} \quad \forall((i,t),(j,u)) \in A_K, \quad \forall k \in K, \tag{17}$$

$$z^p_{(i,t),(j,u),s} \in \{0,1\} \quad \forall((i,t),(j,u)) \in A_P, \quad \forall p \in P, \forall s \in S, \tag{18}$$

$$a_{p,s} \in \{0,1\} \quad \forall p \in P, \forall s \in S. \tag{19}$$

3.3.3 The Deterministic Second-Stage and Single-Stage Model

Although demand is not known at the first stage, the two-stage program allows us to find a vehicle routing schedule that maximises the expected number of fulfilled orders. Based on this schedule $\widehat{\boldsymbol{x}}$, the actual flow of parcels can be planned at the second stage once demand is revealed. For this, we define the deterministic second-stage model for a given vehicle routing $\widehat{\boldsymbol{x}}$ and a realised demand scenario s as follows:

$$\begin{aligned} \max \quad & \sum_{p\in P_s} a_{p,s} \\ \text{s.t.} \quad & (\widehat{\boldsymbol{x}}, z_s, a_s) \in C_s. \end{aligned} \tag{\textit{2nd-}\mathcal{SP}}$$

The feasible set C_s is defined as above, Constraints (1) to (19), with the only difference that $\widehat{\boldsymbol{x}}$ is treated like a given value instead of a decision variable. The objective of Equation (*2nd-$\mathcal{SP}$*) is to serve as many parcel requests as possible.

In Section 4 we present a scenario decomposition heuristic which requires solving the routing of vehicles and the flow of parcels simultaneously for a given realised demand scenario. To that end, we define the deterministic single-stage problem for a specific scenario $s \in S$ similar to Model (*2nd-$\mathcal{SP}$*) as:

$$\begin{aligned} \max \quad & \sum_{p \in P_s} a_{p,s} \\ \text{s.t.} \quad & (x_s, z_s, a_s) \in C_s, \end{aligned} \tag{\textit{1-$\mathcal{SP}$}}$$

where now x_s constitutes a scenario-dependent decision variable.

4 Multiple Scenario Approach

Finding a consistent vehicle routing schedule between the micro-hubs is a difficult problem for many reasons. First, it requires solving a combinatorial optimisation problem on the first stage. It includes solving a VRP, and hence is a $\mathcal{NP}$-hard problem. Second, the first stage solution must respond to the uncertainty of the second stage. The extended two-stage model where all scenarios are considered simultaneously cannot be solved directly within reasonable time, even for small instance sizes. We hence require an heuristic approach to reduce the problem size. Scenario-decomposition techniques are inherent to two-stage stochastic programs which naturally can be decomposed into scenario-dependent problems. Thus, we decide to use a multiple scenario approach to derive a consistent first stage solution. We introduce the general framework of the MSA in Section 4.1. At its core, a consensus functions assesses several scenario-dependent solutions to detect common patterns among individual solutions. The consensus function thus represents the most important component of the algorithm. In Section 4.2, we present different consensus functions analysing different aspects of the individual solutions.

4.1 General Framework

In this section we present the MSA which was introduced by Bent and Van Hentenryck (2004). In Algorithm 1 we state the pseudo code of the algorithm, which we explain in the following. The MSA makes use of a set of several sampled scenarios that represent potential realisations of the uncertain demand, denoted by S. For each scneario $s \in S$ the scenario-dependent objective function and feasible set are denoted by f_s and C_s, respectively. In step 1, the deterministic single-stage model (*1-$\mathcal{SP}$*) is solved to optimality for each of the sampled scenarios (via Gurobi, see Section 4.3), resulting in an objective value (x_s, z_s, a_s) for all $s \in S$. Then, in step 2, each of these individual solutions is evaluated with a consensus function $g(\cdot,\cdot)$ that compares how similar this solution is to any other solution. The choice of the consensus function plays a major role within MSA as it significantly influences the solution the approach produces. For a pair of scenarios s and $\widehat{s}$, the consensus function $g(x_s, x_{\widehat{s}})$ compares the solutions corresponding to the first stage, x_s and $x_{\widehat{s}}$. The MSA assigns a score $\sigma(s)$ to each scenario $s \in S$ which is determined as the sum of all values of the consensus function between s and any other scenario $\widehat{s} \in S \setminus s$. In step 3, the scenario with best score is chosen. The solution of this scenario shows the highest similarity to other scenario-dependent solutions according to the consensus function g. This is the reason why the choice of g mainly characterises the resulting solution. Note that depending on the definition of g, the "best" score may either be the minimum or the maximum score. This will be detailed below when introducing consensus functions. Finally, the solution of the scenario with best score is selected as consistent first stage solution of the two-stage stochastic program (*2-$\mathcal{SP}$*).

4.2 Consensus Functions

In this section we elaborate several consensus functions that aim to detect problem-specific similarities in the scenario-dependent solutions. According to the assessment by the consensus function, the MSA then

Algorithm 1: Multiple Scenario Approach

Input: A set of possible scenarios S with an objective function f_s and a feasible set C_s $\forall s \in S$; and a consensus function $g(\cdot, \cdot)$.

1. $\forall s \in S : (x_s, z_s, a_s) := \text{argmax} \left\{ \sum_{p \in P_s} a_{p,s} \text{ s.t. } (x, z, a) \in C_s \right\}$ (*1-$\mathcal{SP}$*).
2. $\forall s \in S : \sigma(s) := \sum_{\widehat{s} \in S \setminus s} g\left(x_s, x_{\widehat{s}}\right)$.
3. Choose $\bar{s} \in \underset{s \in S}{\text{argmin}} \left\{\sigma(s)\right\}$.

Output: $x_{\bar{s}}$ as solution to the first stage of (*2-$\mathcal{SP}$*).

derives a consistent first stage solution. Recall that in the example above (Section 3.2) both micro-hubs are visited, yet not all parcel orders can be fulfilled. In order to be flexible with respect to uncertain pickup and delivery requests, first stage solutions need to balance various aspects. These are the *order* of visited micro-hubs, *direct connections* between micro-hubs, the *reachability* among micro-hubs, and the *arrival time* at micro-hubs.

micro-hubs should be visited in a meaningful *order* allowing to transport as many of the realised daily parcel requests as possible. Many parcels have to be first picked up at different micro-hubs, and then delivered to other micro-hubs. Hence, it makes a difference in which position of the tour a micro-hub is visited, and which micro-hubs are scheduled before and next. Thus, the order of micro-hubs is an important characteristic of a solution. Not only the entire sequence is of relevance, also *direct connections* between micro-hubs play a role in how many parcel can be delivered. We thus compare direct connections between micro-hubs to detect similarities among solutions. Further, a good first stage solution should be able to transport parcels from any micro-hub to any other micro-hub. Thus, a high *reachability* between micro-hubs is of importance. Moreover, vehicles should arrive at micro-hubs neither too early (to not miss later arriving parcels), nor too late (to respect customer time windows). This is why we consider the *arrival time* at micro-hubs as another comparison criterion. We design one consensus function for each of the mentioned aspects (order, direct connections, reachability, arrival time), presented in Section 4.2.1 to Section 4.2.5. We conclude this section with an illustration of the consensus functions in Section 4.2.6.

4.2.1 Original Consensus Function: Comparing Arcs

The first consensus function makes a full comparison of solutions and is the canonical choice in the MSA as it originates from Bent and Van Hentenryck (2004). It compares two solutions exactly arc per arc in the time extended network. With this, it is straight forward, and neglects any problem specific properties, To do so, the consensus function compares the first-stage decision variables of any two scenarios s and $\widehat{s}$ entry-wise. For ease of notation, we define:

$$\mathscr{A}_K := \left\{((i,t),(j,u),c) \mid (i,t),(j,u) \in A_K,\ c \in K\right\}, \tag{20}$$

and let $a \in \mathscr{A}_K$ represent the indices for vehicle $c \in K$ and vehicle arc $((i,t),(j,u)) \in A_K$. Then, for each arc in the time extended network, the consensus function considers the absolute difference between the two solutions x_s and $x_{\widehat{s}}$:

$$g(x_s, x_{\widehat{s}}) := \sum_{a \in \mathscr{A}_K} \left|x_{a,s} - x_{a,\widehat{s}}\right|. \tag{21}$$

This is an exact comparisons of tours concerning space and time. The consensus function takes value 0 only if the routing in two scenarios is exactly the same. A small deviation in time for example already leads to a high value of the consensus function. Using this consensus function, the scenario solution with minimal score is rated best since this is the one most similar to other scenario solutions. We will refer to this consensus function as `BvH`.

4.2.2 Comparing the Order of Micro-Hubs

Motivated by the idea that two tours that are equal despite a time shift of a few minutes can still be seen as quite similar to each other, we develop other consensus functions being less restrictive. An option to compare similarity of routes in space and time in a softer variant is to use the Hamming distance. The Hamming distance compares the position of micro-hubs in a tour between two solutions. This way, the MSA has an eye on the *order* in which micro-hubs are visited, which can have a large impact on how many parcels can be served. For this, we create an ordered list of the micro-hubs a vehicle visits on its tour. For scenario $s \in S$ let

$$h_s := (h_{s,1}, h_{s,2}, ..., h_{s,n_s}), \tag{22}$$

with $n_s \in \mathbb{N}$ and $h_{s,i} \in V_H \ \forall i \in [0, n_s]$ denote the order of visited micro-hubs. Note that n_s might be different for different $s \in S$. Further, as some micro-hubs might be visited several times, it is possible that $h_{s,i} = h_{s,j}$ for $i, j \in [0, n_s], i < j$. With this, we define the consensus function based on the Hamming distance as:

$$g(x_s, x_{\widehat{s}}) := \sum_{i=0}^{\min(n_s, n_{\widehat{s}})} \mathbb{1}_{\{h_{s,i} = h_{\widehat{s},i}\}}. \tag{23}$$

By considering the order of visited micro-hubs, we still compare routes quite tightly in time and space. However, we drop the strict comparison of arrival times at micro-hubs and call two scenario solutions already "identical", if the order of visited hubs coincides. With the Hamming-distance-based consensus function, the scenario with highest score is most similar to all other ones and hence rated best. We will refer to this consensus function as `order`.

4.2.3 Comparing Direct Connections

Next, we propose a consensus function that focuses on similarity in space. For this, we consider tuples of micro-hubs that are *directly connected*, i.e. that are neighboured in h_s meaning that are visited directly after each other. This reflects that for a consistent first stage solution it is relevant "from where to where" vehicles continue their route. To compare a scenario s to another scenario $\widehat{s}$, we check for each tupel $(h_{s,i}, h_{s,i+1})$ in h_s with $i \in [0, n_s - 1]$ if it is also present in $h_{\widehat{s}}$. Formally, we define the corresponding consensus function as:

$$g(x_s, x_{\widehat{s}}) := \sum_{i=1}^{n_s} \mathbb{1}_{\{(h_{s,i}, h_{s,i+1}) \subseteq h_{\widehat{s}}\}}. \tag{24}$$

To assess a scenario we hence check whether each direct connection between two micro-hubs is also present in the other scenarios. Thereby, it does not matter at what time the vehicle traverses between the two micro-hubs. Thus, we concentrate on evaluating direct micro-hub connections (i.e. blocks of two micro-hubs) regardless of the arrival times. Two scenarios are evaluated as more similar, the more direct hub connections they share. Therefore, the scenario with highest score is rated best. We will refer to this consensus function as `direct`.

4.2.4 Comparing Reachability

The next consensus function we suggest softens the similarity in space to allow even more flexibility in schedules as Equation (24). Instead of comparing direct connections between micro-hubs, we are now interested in the *reachability* among micro-hubs. For this, we check if there is a path between two micro-hubs at all - possibly with intermediate stops at other micro-hubs. For two scenarios, we compare if a micro-hub can be reached from a certain different micro-hub. As mentioned above, the existence of a path from micro-hub A to another micro-hub B largly influences whether parcels can be transported from A to B at all. To this end, we define the reachability matrix $R_s \in \{0,1\}^{|V_H| \times |V_H|}$ for scenario s with entries:

$$R(i,j)_s := \begin{cases} 1, & \text{if there is a path from hub } i \text{ to hub } j \text{ in solution } x_s, \\ 0, & \text{otherwise.} \end{cases} \tag{25}$$

This allows us to compare whether micro-hub j is visited at some time after micro-hub i. It is possible, that other micro-hubs are visited in between. In other words, if $R(i,j)_s = 1$, then micro-hub j can be reached from micro-hub i in scenario s. To formalise this, we define the consensus function comparing paths between hubs as:

$$g(x_s, x_{\hat{s}}) := \sum_{i \in V_H} \sum_{j \in V_H} |R(i,j)_s - R(i,j)_{\hat{s}}| \,. \tag{26}$$

The consensus function thus checks for two scenarios if there is a path from one micro-hub to another for all pairs of micro-hubs. That way, two solutions are still compared with respect to time and space, but in a much softer variant as in Equations (21), (23) and (24). As in Section 4.2.1, the scenario solution with minimal score is most similar to the other ones with respect to the reachability matrix. We will refer to this consensus function as `reachability`.

4.2.5 Comparing Arrival Times

Lastly, we propose another consensus function that is of different nature than the previous ones. While Equations (23) and (24) and also Equation (26) focus more on the spacial structure of a solution by relaxing the comparison of arrival times, we now consider the *arrival time* at micro-hubs only. This is motivated by the fact, that micro-hubs should neither be visited too early nor too late to transport as many parcels as possible despite their release times and delivery deadlines. More precisely, we compare which micro-hubs are visited in the same hour of the service time window. To formalise this, we define

$$\mathscr{H}_s(t) := \{h_{s,i} \in h_s \mid \text{ arrival time at micro-hub } h_{s,i} \text{ lies in } [t, t+60]\}, \tag{27}$$

as the set of micro-hubs that is visited in hour $[t, t+60]$ in scenario s. Then, as consensus function between two scenarios, we count the number of micro-hubs that are visited within the same hour using the following definition:

$$g(x_s, x_{\hat{s}}) := \sum_{t \in \{0,60,\dots,T_{max}-60\}} \sum_{h \in \mathscr{H}_s(t)} \mathbb{1}_{\{h \in \mathscr{H}_{\hat{s}}(t)\}}. \tag{28}$$

This consensus function ignores any spacial patterns of the solution but aims to detect whether certain micro-hubs are often visited early or late, respectively. Due to its definition, the scenario with highest score is chosen to be the most similar one. We will refer to this consensus function as `time`.

4.2.6 Example

We conclude this section with a short example illustrating the different consensus functions. We consider two scenarios, Table 3 shows the order of visited micro-hubs as well as corresponding arrival times at the micro-hubs. As the arc-based consensus function compares each arc in the time-extended network, it is too large to display. For the remaining consensus functions, Table 4 illustrates the different concepts.

scenario 1	visited micro-hubs (h_1)	1	2	3	4
	arrival time	20	40	65	85
scenario 2	visited micro-hubs (h_2)	2	3	1	4
	arrival time	30	55	85	110

Table 3: Order of visited micro-hubs and corresponding arrival times in two exemplary scenarios.

The original consensus function `BvH` takes value 0 since no arc is used in common in the two scenarios. Only micro-hub 4 is at the same position in the two scenarios. Thus, consensus function `order` takes value 1. Consensus function `direct` takes value 1 since there is one common direct connection between micro-hubs. In both scenarios, the vehicle visits micro-hub 3 after micro-hub 2. Consensus function `paths` compares for which pair of micro-hubs there is a path in the solution, and takes value 4 in the example. Here, four paths are shared: in both scenarios, the vehicle reaches micro-hub 4 from icro hubs

Consensus Function	Scenario	Illustration	$g(x_1, x_2)$
`BvH`	1 2	$\|x_1 - x_2\|_1$	0
`order`	1 2	1 2 3 $\underline{4}$ 2 3 1 $\underline{4}$	1
`direct`	1 2	1 $\boxed{2\ \ 3}$ 4 $\boxed{2\ \ 3}$ 1 4	1
`reachability`	1 2	1 2 3 4 2 3 1 4	4
`time`	1 2	1^{st} hour: 1, 2 2^{nd} hour: 3, 4 1^{st} hour: 2, 3 2^{nd} hour: 1, 4	2

Table 4: Illustration of different consensus functions.

1, 2, and 3, and micro-hub 3 from micro-hub 2. Consesnus function `time` checks which micro-hubs are visited within the first, and second hour. micro-hubs 2 and 4 are visited within the same interval, hence the consensus function takes value 2.

4.3 Implementation

In this section we explain how the different consensus functions are evaluated in our computational study. For each consensus function, we choose the best solution among ten trials. For this, we proceed as follows. First, we generate a set of 20 training scenarios (see Section 5.1). Each of these scenarios is solved using model (*1-$\mathcal{SP}$*) implemented in Gurobi 9.1. (Gurobi Optimization (2021)). For optimisation in Gurobi, we set a time limit of 2 hours. All instances (but very large ones, see Section 5.3.2 for details) are solved to optimality within this time limit. We further provide a feasible starting solution in which the van stays in the depot, and hence no parcels are transported. Then, the different consensus functions are applied to determine the corresponding MSA solutions. This procedure is repeated ten times, we thus derive ten MSA solutions for each consensus functions. They are evaluated on a new set of 20 test scenarios using the deterministic second-stage model (*2nd-$\mathcal{SP}$*), and average objective values are computed. We then choose the best MSA solution as the one with highest average objective value. This way, one "best" MSA solution is derived for each consensus function. For final evaluation, these best MSA solutions are assessed on 100 new generated scenarios with the deterministic second-stage model (*2nd-$\mathcal{SP}$*), and objective values are averaged per consensus function.

5 Computational Study

In this section we present the experimental setup and results of our computational study. In Section 5.1, we explain how instances are generated. In Section 5.2, we introduce benchmark solutions to assess the solutions obtained by our approach. In Section 5.3, present our computational results.

5.1 Instance Generation

The following explains how we generate different possible demand scenarios. We assume demand to be distributed over a square city area with a radius of $r = 10$ (km). We motivate our instances by the city structure of Braunschweig, Germany, see (Ulmer and Streng 2019). Braunschweig shows the classical European city structure with a city centre and several ring roads. Several parcel locker stations of the German post service DHL are placed on the main ring road in Braunschweig. Inspired by this, we place 5 micro-hubs equidistantly on a circle with a radius of 5 (km), which is half of the city radius. Moreover, with such a circular structure the micro-hubs are evenly spread over the city area. The depot is located at the middle of the upper edge of town. Figure 4 gives an illustration of the circular location of five micro-hubs and the depot in a square city. We assume that each micro-hub has a maximum capacity of 20 parcels. One delivery vehicle conducts service between micro-hubs. In our computational study, we assume the vehicle to be a large cargo bike with a speed of 25 ($^{\mathrm{km}}/_{\mathrm{h}}$) and maximum capacity of 20 parcels.

For our experiments, we test different delivery promises and service time horizons. We investigate three service designs: "instant" – instant delivery (60 min.) in a short horizon (240 min.); "fast" – fast delivery (120 min.) in a medium horizon (360 min.), and "same-day" – delivery on the same day (480 min.) in a large horizon (480 min.). The latter is equivalent to not imposing customer time windows. We use a discrete step size of $\delta := 10$ minutes in the time expanded network. The travel time $t_{i,j}$ between two locations $i, j \in V_H \cap V_0$ is computed via the Euclidean distance between i and j divided by the vehicle's velocity, and is then rounded up to the next multiple of δ.

For each service design, we run our experiments with a varying number of parcels, $|P| \in \{80, 100, 120, 140\}$. For each parcel request we sample a release time, a pickup store and a delivery customer location within the city area. All parcels have a homogeneous volume of one. The release time of a parcel is drawn uniformly over the time horizon $T = \{0, 10, ..., T_{max} - 120\}$.

Pickup micro-hub (origin) and delivery micro-hub (destination) of a parcel are the micro-hub closest to the corresponding pickup store and delivery customer, respectively. For spatial distribution of stores and customers we consider two different demand patterns:

- *uniform*: Stores and customers are uniformly distributed over the entire city area. An example of this customer distribution is shown on the right-hand side of Figure 4. This is inspired by the city structure of Göttingen, Germany, where stores can be found over the entire city area, and inhabitants live both inside and outside he city centre.
- *clustered*: Stores and customers are clustered within the city. Inspired by the city structure of Braunschweig, we designate the inner part of the city as city centre, the south-western part as industrial area, and northern as well as eastern part as residential area. The exact layout is shown on the left-hand side of Figure 4. Stores are located in the city centre and industrial area; customers mostly in residential, but also in the industrial area. More details are presented in Appendix A.1.

In Table 5 we summarise the parameter values used for scenario generation in the computational experiments.

5.2 Benchmarks

We compare the solutions obtained by the MSA approach with the consensus function proposed in Section 4 to two benchmark solutions. First, we use a solution in which routing consistency constraints are relaxed. To this end, we solve the deterministic single-stage model (*1-$\mathcal{SP}$*) separately for each scenario. As this is done scenario-dependent, we can determine vehicle routing and parcel flow simultaneously on one stage. Given optimal solution for this "daily" problem, this constitutes an upper bound to the MSA. Since this solution requires replanning on a daily basis, we refer to this as the daily solution.

Second, we implement a practically-inspired fixed solution that aims on a high flexibility and reachability among micro-hubs. The solution should allow to reach all micro-hubs from any micro-hub, and further should not loose too many time between any two visits. Therefore, we suggest a circular route. The vehicle leaves the depot and visits each micro-hub, in ascending order. When the last micro-hub is

Description	Notation	Instant	Fast	Same-Day
service time horizon	T_{max}	240	360	480
delivery promise	T_P	60	120	480
nr. parcels	$\|P\|$	80, 100, 120, 140		
demand pattern	–	*uniform*, *clustered*		
nr. micro-hubs	$\|V_H\|$	5		
nr. vehicles	$\|K\|$	1		
max. capacity micro-hub	C_H	20		
max. capacity vehicle	C_K	20		
vehicle velocity	v_K	25		
parcel volume	u_p	1		
length of discrete time step	δ	10		

Table 5: Parameter values used in computational experiments.

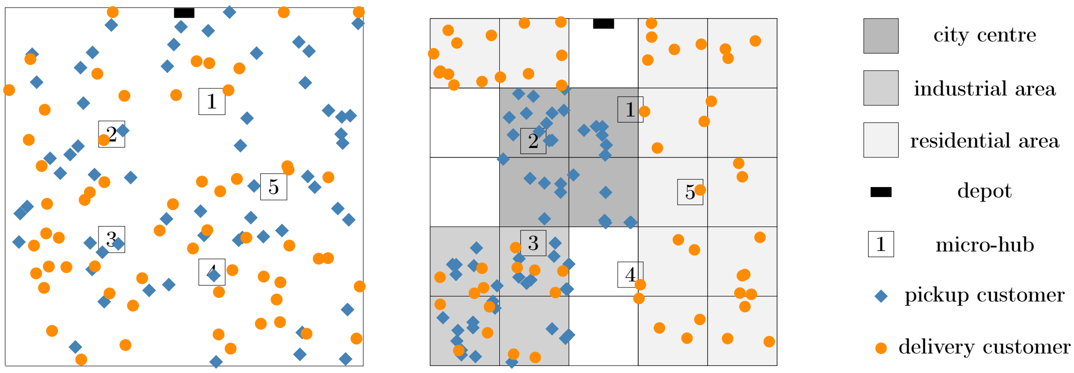
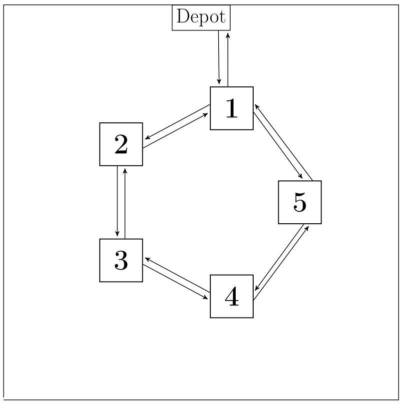


Figure 4: Example of demand structure *uniform* (l.h.s.) and *clustered* (r.h.s.) with 60 parcel orders.

reached, the vehicle continues to micro-hub 1 to close the circle. Instead of traversing the same circle again, the vehicle turns and takes the same route back to the depot, visiting all micro-hubs again, but in descending order. This guarantees that not too much time is needed for delivery between two micro-hubs. If the service time horizon is longer than the total tour length, the vehicle starts its tour such that the tour finishes at T_{max}. This is motivated by the idea to increase consolidation opportunities: visiting micro-hubs at a later point in time may allow to transport more parcels, as more demand will arise during the course of the day. In Figure 5, we provide an illustration of the fixed solution for $T_{max} = 360$.

5.3 Computational Results

We present our computational results in the following. In Section 5.3.1, we conduct a method analysis, investigating the effect of different service designs and demand patterns. We further assess the value of micro-hubs as well as the cost of consistent routing, and analyse two exemplary first-stage schedules. In Section 5.3.2, we analyse the solution method by looking at the different consensus functions in more detail, and shedding a light on very large instances.



Location	Arrival	Departure
depot	—	120
micro-hub 1	140	150
micro-hub 2	160	170
micro-hub 3	180	190
micro-hub 4	200	210
micro-hub 5	220	230
micro-hub 1	240	250
micro-hub 5	260	270
micro-hub 4	280	290
micro-hub 3	300	310
micro-hub 2	320	320
micro-hub 1	340	330
depot	360	—

Figure 5: Fixed solution for $T_{max} = 360$, in minutes after the start of the service horizon.

5.3.1 Problem Analysis

We investigate the average service rates that are obtained by the MSA and benchmarks solutions, and derive various insights from this. For each service design (instant, fast, same-day) we compare our best performing consensus function to the original one from Bent and Van Hentenryck (2004) (`BvH`), as well as the two benchmark solutions. Figure 6 shows the results for *uniform* demand on the left-hand side, and for *clustered* on the right-hand side. In the following, we look at the impact of the different delivery time promises and the different (spatial) demand patterns. We further analyse the value of consolidation at micro-hubs and evaluate the cost of routing consistency.

Service Designs. With an instant delivery service design, with our approach about 36.79% of all parcel orders can be delivered on *uniform* instances and 36.74% on *clustered* instances. However, the daily, upper bound, solution also permits to serve only 43.91% (on *uniform*) and 41.17% (*clustered*) of all parcel orders on average, respectively.

In the service design of fast delivery, more time is available, resulting in approximately 74.20% parcels being served with our approach on *uniform* demand, and 51.06% parcels on *clustered* demand. This clearly outperforms both the fixed solution, but also the original consensus function for both demand patterns. Because of the longer time horizon, the number of arcs in the time expanded network is much higher than in the instant service design. Thus, it is less likely for the original consensus function to detect common arcs. This is the reason why it performs significantly worse than our consensus functions compared to instant delivery (in relative sense). The fixed solution is very restrictive and lacks flexibility. Due to the limited delivery time promise of two hours, it hence does not allow to serve many parcel orders in a fast service design.

In the service design of same-day delivery, all solutions perform well because of the relatively high flexibility in time. We reach nearly day-optimal solutions with our approach: with our consensus functions we are able to serve 83.90% of all parcel orders on *uniform* demand, and 56.18% of all parcel orders on *clustered* demand. For both demand patterns, this is very close to the daily solution: with our solution 0.31% parcels are delivered less on *uniform* demand, and 0.37% parcels less on *clustered* than with the daily, upper bound solution.

To summarise, the average service rates obtained by our approach are higher than the ones from the original consensus function and benchmark solutions for any service design. Average service rates are generally higher for service designs with softer time constraints due to increased flexibility. From this, we deduce that the more time is available, the better MSA performs. With same-day delivery the MSA reaches objective values being nearly day-optimal.

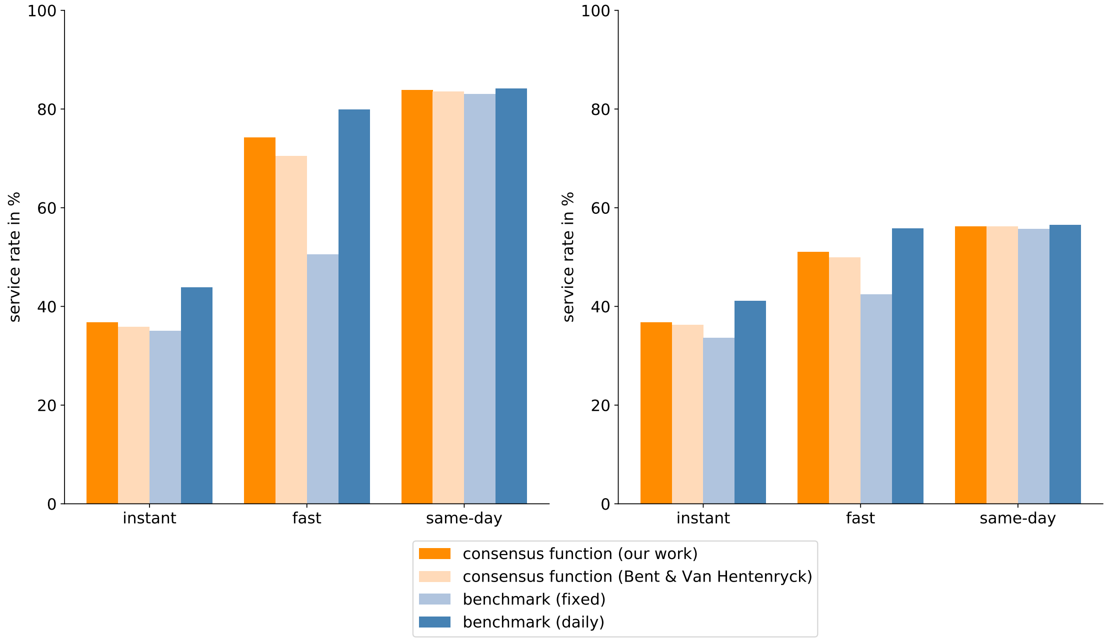


Figure 6: Average service rates on *uniform* (l.h.s.) and *clustered* (r.h.s.) demand.

Demand Patterns. While our approach outperforms benchmark solutions and the original consensus function for any service design and both demand patterns, we observe difference for *uniform* and *clustered* demand in Figure 6. The average amount of parcels that can be delivered on *uniform* instances is much higher than on *clustered* instances, especially in a fast and same-day service design. In the clustered demand setting, some micro-hubs show to have more pickup requests, while others mainly serve as drop-off location. We hence observe a repeated sequence of picking up and dropping off parcels (see Section 5.3.1 for more details). This indicates, that the vehicle has a low fill rate when returning from a drop-off micro-hub to pick up new parcels at a different micro-hub. Also, due to this structure, more time is spent on the street. This explains the lower service rates for *clustered* compared to *uniform* demand.

However, the average gap of our solution to the daily one is smaller for *clustered* demand. With instant delivery, the MSA deviates by 18.03% from the daily solution for *uniform* demand, while the gap is only 12.13% for *clustered* demand. With same-day delivery, the MSA deviates by 0.79% from the daily solution on *clustered* demand, which is about a third less than on *uniform* demand. We hence conclude that the MSA yields results closer to the daily upper bound if more structure of the underlying demand scenario is known.

The Value of Micro-Hubs. We also analyse the value of micro-hubs, i.e. the value of consolidation in a local transportation system. In Figure 6 we see that average service rates are quite low for an instant delivery service, even for the upper bound solution. This indicates that it is difficult to use micro-hubs for serving a high number of parcels within a short delivery time promise, even without consistent routes. For such business models, the gain in consolidation is limited due to narrow delivery deadlines and direct transportation may be the more suitable choice. This changes for fast and same-day delivery: due to more temporal flexibility, high amounts of parcels can be delivered when using the MSA or the daily solution. In contrast, the fixed solution leads to a very high loss in service quality. Using the right strategy, consolidation at micro-hubs is hence worthwhile for service designs with longer delivery time promises, especially on *uniform* demand.

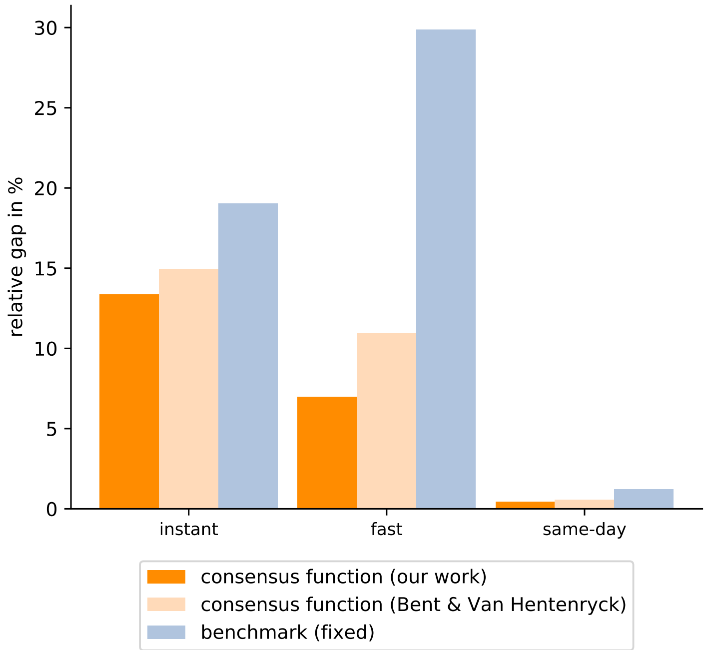


Figure 7: Average gap to daily solution.

The Cost of Consistency. In this section we investigate the cost of consistency of our solution as well as the original consensus function and the fixed benchmark. For this, we assess the average relative gap of these solutions to the daily solution, where consistency constraints are omitted. To this end, we compute the relative gap of the MSA solutions as well as the fixed solution to the daily solution. Let z_{daily} and z denote the objective value of the daily and the MSA/fixed solution, respectively. Then the relative gap of the MSA/fixed solution to the daily solution is

$$\frac{z_{daily} - z}{z_{daily}}. \tag{29}$$

In Figure 7 we display the relative gaps of our best consensus function, the consensus function `BvH`, and the fixed solution, averaged over all instances. It is striking, that our approach clearly outperforms both the original consensus function `BvH` and the fixed solution on all instances on average. With an instant service design, the first-stage solution produced by our approach deviates by about 13.38% from the daily solution. In contrast, consensus function `BvH` deviates by 14.95%, and the fixed solution by 19.05%. For fast delivery, the average relative gap of our approach can be reduced to 6.98%. With a same-day delivery promise, the solution found by our approach allows to deliver only 0.44% parcels less than with the daily solution. Thus, consistency in tours comes to some cost for instant and fast service designs. For same-day delivery however, consistent routes can be implemented without loosing much service quality compared to a non-consistent daily re-optimised routing. With the right strategy, it is therefore possible to implement consistency at very low "cost".

Routing Structure. To conclude our computational analysis, we examine two exemplary routes in more detail. We consider instances with 120 parcels and the same-day service design, and investigate those consensus functions performing best for these instance classes. This is `order` for *uniform* demand, and `BvH` for *clustered* demand, for details see Appendix A.2, Figure 10. Note that `BvH` performs better than other consensus functions for this instance as an exception, see Section 5.3.2. The resulting routes are displayed in Figure 8. micro-hubs are represented by two half circles. The left-hand side illustrates

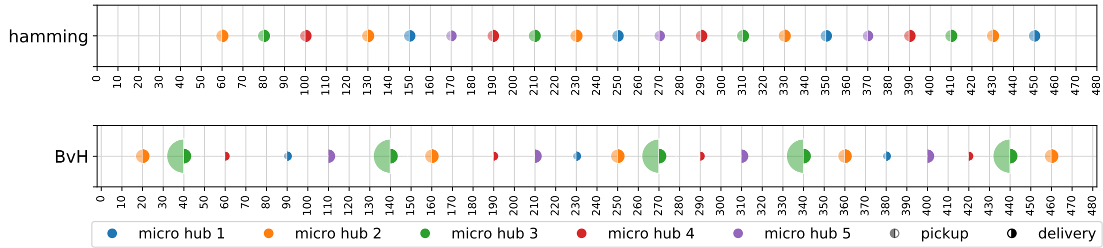


Figure 8: Routes produced by the MSA with consensus function `order` and `BvH` for large instances with 120 parcels and *uniform* and *clustered* demand, respectively.

the relative amount of pickup requests that originates from that specific micro-hub on average over 200 instances. Analogously, the right-hand side illustrates the relative average amount of delivery requests destined to this micro-hub.

The upper part of the figure shows the route produced by the MSA with consensus function `order` for *uniform* demand. First, we see that – as demand is uniformly distributed – all micro-hubs have approximately the same amount of parcels that have to be picked up there, and delivered to this destination. Second, we see that the vehicle waits at the depot at the beginning of the service time horizon and visits the first micro-hub at minute 60. In the beginning of the service time horizon, not many parcel requests are present, and additionally they are distributed uniformly over all micro-hubs. It hence becomes meaningful to wait some time at the beginning until a sufficient amount of parcel requests has been placed to make it worthwhile to visit a micro-hub. The vehicle first visits micro-hubs 2, 3, and 4. Then, it goes back to micro-hub 2 and further conducts a circular route: it travels a clockwise circle three times before going back to the depot. Since pickup and delivery locations are uniformly spread, such a circular structure provides a time-efficient route visiting many micro-hubs, and thus allowing to deliver many of the unknown parcel requests.

The lower part of Figure 8 shows the route produced by the MSA with consensus function `BvH` for *clustered* demand. We clearly see the structured demand pattern: Most parcels originate from micro-hub 3, some from micro-hub 2 and even fewer from micro-hub 1. Those micro-hubs are located in the city centre or industrial zone and hence are mainly used as pickup locations, cf. Figure 4. micro-hubs 4 and 5 are located within or close to residential areas, and consequently barely have pickup requests (4.24% and 2.28% of all parcel requests, respectively). Parcels' destinations are spread more evenly among micro-hubs. Most parcels are destined to micro-hubs 2, 3, or 5 respectively, serving the industrial zone, the north-western and eastern residential areas. micro-hubs 1 and 4 have slightly less delivery requests, serving the north- and south-eastern residential area. The route produced by the MSA with consensus function `BvH` captures this demand pattern: the vehicle first visits micro-hubs with most pickup demand (micro-hubs 2 and 3) and then delivers the collected parcels to their destinations. After that, it comes back to micro-hubs 3 and 2, to collect more parcels, which are again delivered. This process of collection and delivery of parcels is repeated until the service time horizon ends.

This example shows that the MSA is able to detect common patterns in the underlying demand scenarios. It utilizes this to derive a consistent first stage solution that adapts to this structure. It thus provides a powerful tool to develop a consistent route between micro-hubs. The method is even more powerful if some underlying demand pattern is available since it then can make use of these structural similarities.

5.3.2 Method Analysis

The Value of Our Consensus Functions. Next, we discuss the performance of the different consensus functions proposed in Section 4.2 in more detail. To get a better understanding of the different consensus functions, we calculate the average relative gap to the daily solution for each consensus function. The results are depicted in Figure 9. We further count how often each consensus function

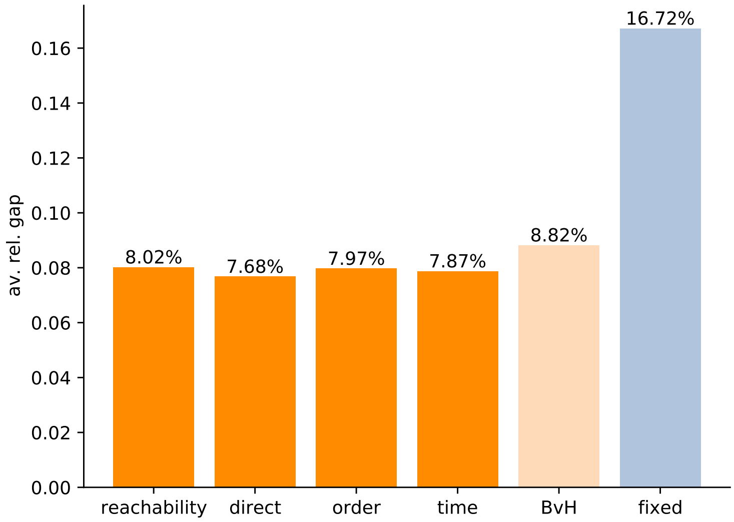


Figure 9: Average relative gap to daily solution per consensus function.

`reachability`	`direct`	`order`	`time`	`BvH`
6	9	9	7	2

Table 6: Count of how often which consensus function yields the best objective value.

yields the best objective value, depicted in Table 6. Note that on some instances several consensus functions yield the same best objective value.

We see that the MSA clearly outperforms the fixed solution and the original consensus function, regardless of the consensus function choice. The fixed solution has a gap of 16.72% to the daily solution on average, the original consensus function a gap of 8.82%. Although the different consensus functions uncover different structures in the solutions, they all yield similar objective values and gaps, which vary between 7.68% and 8.02%. Thus, there is no unique best consensus function, each of them yields the best solution between 6 to 9 times. Our consensus function outperform the original one `BvH` in 91.66% of the investigated cases. On very few instances only, the original consensus function `BvH` performs better. We derive the following insights from this. First, the MSA is an effective method to derive consistent master tours in a stochastic environment compared to a fixed solution. Using MSA reduces the gap to the daily-optimal solution by 9.04% compared to the fixed circular route. Second, making use of problem-specific properties in the consensus function increases the effectiveness of the method. On average, the novel consensus functions we propose perform up to 1.54% better than the original consensus function suggested by Bent and Van Hentenryck (2004).

Suboptimality on Large Instances. We observed that the model Equation ($\mathit{1}$-$\mathcal{SP}$) can not be solved to optimality by Gurobi within the given time limit of 2 hours for instances with *uniform* demand and 140 parcels If Gurobi did not find the optimal solution, the current best (possibly sub-optimal) solution was used for processing the MSA solutions. Out of 200 training scenarios, only 43 were solved to optimality within the given time limit. For 157 of the 200 training scenarios, Gurobi returned a MIP gap of infinity, implying that either no objective bound is available, or no other solution different from

av. obj. val.	`reachability`	`direct`	`order`	`time`	`BvH`	`fixed`	`daily`
absolute	99.06	99.06	99.06	98.92	98.72	98.76	96.79
relative	70.76%	70.76%	70.76%	70.66%	70.51%	70.54%	69.14%

Table 7: Average absolute and relative amount of delivered parcels for large *uniform* distributed instances with 140 parcels.

the starting solution (which has an objective value of 0) has been found. The remaining 5 scenarios have a relative MIP optimality gap of 3.14%.

Although the single-stage model ($1\text{-}\mathcal{SP}$) cannot be solved to optimality, the MSA – which takes these sub-optimally solved instances as basis – yields better results than the daily solution, as can be seen in Table 7. With the consensus functions `reachability`, `direct`, and `order`, the MSA leads to solutions that are 2.38% better than the daily one on average. Note that the daily solution solves the single-stage model separately for each scenario, hence the same problem of potential sub-optimality occurs. This observation is striking: the MSA does not necessarily need optimally solved scenario-dependent solutions to produce a consistent first stage solution that performs well – and even better than the daily one – under different demand scenarios. In contrary, optimal scenario-dependent solutions might be "overfitted" to a specific scenario, aiming to maximise the amount of delivered parcels on that specific scenario. A non-optimal solution might be more flexible: although suboptimal in the specific scenario, it may allow the transportation of more (unknown) parcels in a distinct scenario instead.

6 Conclusion

In this paper, we have analysed the value and functionality of consistent collaborative delivery with micro-hubs for a local market. We have proposed a novel model formulation, a two-stage stochastic integer program where on the first stage tours are determined and on the second stage parcels flows are optimised. We have solved the problem using a multiple scenario approach. We have proposed four new consensus functions that are designed to detect problem-specific similarities among scenario-dependent solutions. The performance of the different consensus functions, the traditional consensus function and two benchmark heuristics have been evaluated in a computational study. Our results show when and how consistent tours for local same-day delivery can be beneficial.

There are several avenues for future research. We have found that with any of the investigated consensus functions the MSA outperforms the fixed benchmark heuristic as well as the original consensus function. However, there is not a unique best candidate, rather different consensus functions perform best for different instance types and sizes. Future work may merge this to an "aggregated" consensus function that either combines the consensus values or adaptively uses the individual ones. We have seen that micro-hubs can be very valuable for deliveries within 2 hours. Future work may further investigate what deliveries are suitable for consolidated shipping and which should be shipped directly. In our research, we have focused on the single-vehicle case for moderately sized cities to analyse the functionality of the MSA and the impact of consistent routes. Future work may extend the computational study to larger cities and fleets. While the proposed model is already designed to capture multiple vehicle, it is very likely that the second stage problems cannot be solved with standard methodology. Instead, metaheuristics might be developed. Another extension could be a combination of the consistent delivery with courier bikes, e.g., for delivering from shops to micro-hubs or from micro-hubs to customers as well as delivering orders that cannot be served within the micro-hub framework. This would lead to a third decision dimension about the routing of the courier bikes. Further, in this research we have assumed that all parcels become known at once during the day. Future work may model the arrival of parcels dynamically every day. This would replace the second stage of our model with a stochastic dynamic process. Determining the consistent tour from a dynamic routing policy of the vehicles would be another, currently unexplored, research opportunity. Furthermore, we have shown that our MSA allows to derive a consistent tour performing not much worse than a daily re-optimised one, and additionally shows the advantage of computing the master tour only once. With this, the daily operational planning can be reduced to the flow of parcels, and the costly routing of the full system can be omitted. For larger scale instances, the decomposition of first and second stage actually leads to better performance compared to a joint optimization without consistency. Thus, determining consistent subparts of a solution may be advantageous even in cases where no consistency is required. Future work may analyse related stochastic problems and how predetermining parts of the solution may improve overall performance. Another interesting observation is that even though some scenarios of the MSA were not solved to optimality for the larger size instances, the MSA still provided superior results. Thus, future work may analyse in general the value (and necessity) of optimal scenario solutions in MSAs.

Acknowledgments

The authors' research is partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Emmy Noether Programme, project 444657906. We gratefully acknowledge their support.

Bibliography

[Angelelli et al. 2017] Angelelli, Enrico ; Archetti, Claudia ; Filippi, Carlo ; Vindigni, Michele: The probabilistic orienteering problem. In: *Computers & Operations Research* 81 (2017), p. 269–281.

[Azi et al. 2012] Azi, Nabila ; Gendreau, Michel ; Potvin, Jean-Yves: A dynamic vehicle routing problem with multiple delivery routes. In: *Annals of Operations Research* 199 (2012), Nr. 1, p. 103–112.

[Bent and Van Hentenryck 2004] Bent, Russell W. ; Van Hentenryck, Pascal: Scenario-based planning for partially dynamic vehicle routing with stochastic customers. In: *Operations Research* 52 (2004), Nr. 6, p. 977–987.

[Bernardo and Pannek 2018] Bernardo, Marcella ; Pannek, Jürgen: Robust solution approach for the dynamic and stochastic vehicle routing problem. In: *Journal of Advanced Transportation* 2018 (2018).

[Bertsimas et al. 1990] Bertsimas, Dimitris J. ; Jaillet, Patrick ; Odoni, Amedeo R.: A priori optimization. In: *Operations Research* 38 (1990), Nr. 6, p. 1019–1033.

[Bundesverband Paket und Expresslogistik e.V. (BIEK) 2021] Bundesverband Paket und Expresslogistik e.V. (BIEK), KE-CONSULT Kurte & Esser GbR: *Möglichmacher in bewegten Zeiten, KEP-Studie 2021 – Analyse des Marktes in Deutschland.* `https://www.biek.de/publikationen/studien.html`. 2021. – Online; accessed 25 February 2022.

[Burns et al. 2022] Burns, Tiffany ; Davis, Andrew ; Harris, Tyler ; Kuzmanovic, Alexandra: *Beyond the distribution center.* `https://www.mckinsey.com/industries/retail/our-insights/beyond-the-distribution-center`. 2022. – Online; accessed 24. June 2022.

[Cameron 2022] Cameron, Isabel: *Metapack integrates with Homerr to offer sustainable deliveries for retailers.* `https://www.chargedretail.co.uk/2022/05/27/metapack-integrates-with-homerr-to-offer-sustainable-deliveries-for-retailers/`. 2022. – Online; accessed 24. June 2022.

[Campbell and Thomas] Campbell, Ann M. ; Thomas, Barrett W.: Challenges and advances in a priori routing. In: *The vehicle routing problem: latest advances and new challenges.* Springer (2008), p. 123–142.

[Crainic et al. 2016] Crainic, Teodor G. ; Errico, Fausto ; Rei, Walter ; Ricciardi, Nicoletta: Modeling demand uncertainty in two-tier city logistics tactical planning. In: *Transportation Science* 50 (2016), Nr. 2, p. 559–578.

[Cuda et al. 2015] Cuda, Rosario ; Guastaroba, Gianfranco ; Speranza, Maria G.: A survey on two-echelon routing problems. In: *Computers & Operations Research* 55 (2015), p. 185–199.

[Dalmeijer and Spliet 2018] Dalmeijer, Kevin ; Spliet, Remy: A branch-and-cut algorithm for the time window assignment vehicle routing problem. In: *Computers & Operations Research* 89 (2018), p. 140–152.

[Datex 2021] Datex: *2021 Update: E-Commerce, Last Mile Delivery and 3PLs*. `https://www.datexcorp.com/2021-update-e-commerce-last-mile-delivery-and-3pls/`. 2021. – Online; accessed 29 June 2022.

[Emadikhiav et al. 2020] Emadikhiav, Mohsen ; Bergman, David ; Day, Robert: Consistent routing and scheduling with simultaneous pickups and deliveries. In: *Production and Operations Management* 29 (2020), Nr. 8, p. 1937–1955.

[Florio et al. 2022] Florio, Alexandre M. ; Feillet, Dominique ; Poggi, Marcus ; Vidal, Thibaut: Vehicle routing with stochastic demands and partial reoptimization. In: *arXiv preprint arXiv:2201.08866* (2022).

[Ghiani et al. 2012] Ghiani, Gianpaolo ; Manni, Emanuele ; Thomas, Barrett W.: A comparison of anticipatory algorithms for the dynamic and stochastic traveling salesman problem. In: *Transportation Science* 46 (2012), Nr. 3, p. 374–387.

[Gurobi Optimization 2021] Gurobi Optimization, LLC: *Gurobi Optimizer Reference Manual*. 2021. – URL `http://www.gurobi.com`.

[Hvattum et al. 2006] Hvattum, Lars M. ; Løkketangen, Arne ; Laporte, Gilbert: Solving a dynamic and stochastic vehicle routing problem with a sample scenario hedging heuristic. In: *Transportation Science* 40 (2006), Nr. 4, p. 421–438.

[Jiang and Li 2021] Jiang, Dapei ; Li, Xiangyong: Order fulfilment problem with time windows and synchronisation arising in the online retailing. In: *International Journal of Production Research* 59 (2021), Nr. 4, p. 1187–1215.

[Koç et al. 2020] Koç, Çağrı ; Laporte, Gilbert ; Tükenmez, İlknur: A review of vehicle routing with simultaneous pickup and delivery. In: *Computers & Operations Research* 122 (2020), p. 104987.

[Kovacs et al. 2014] Kovacs, Attila A. ; Golden, Bruce L. ; Hartl, Richard F. ; Parragh, Sophie N.: Vehicle routing problems in which consistency considerations are important: A survey. In: *Networks* 64 (2014), Nr. 3, p. 192–213.

[Lagos et al. 2019] Lagos, Felipe ; Klapp, Mathias A. ; Toriello, Alejandro: Branch-and-Price for Routing with Probabilistic Customers / Working paper. 2019. – Research Report.

[Metapack 2022] Metapack: *Ecommerce Delivery Benchmark Report 2022*. `https://info.metapack.com/ecommerce-delivery-benchmark-report-2022.html?utm_source=press&utm_medium=referral&utm_campaign=delivery+benchmark+2022`. 2022. – Online; accessed 29 June 2022.

[Neumann-Saavedra et al. 2016] Neumann-Saavedra, Bruno A. ; Crainic, Teodor G. ; Gendron, Bernard ; Mattfeld, Dirk C. ; Römer, Michael: Service Network Design of Bike Sharing Systems with Resource Constraints. In: *CIRRELT-2016-33* (2016).

[Orenstein and Raviv 2022] Orenstein, Ido ; Raviv, Tal: Parcel delivery using the hyperconnected service network. In: *Transportation Research Part E: Logistics and Transportation Review* 161 (2022), p. 102716.

[Ritzinger et al. 2016] Ritzinger, Ulrike ; Puchinger, Jakob ; Hartl, Richard F.: A survey on dynamic and stochastic vehicle routing problems. In: *International Journal of Production Research* 54 (2016), Nr. 1, p. 215–231.

[Salavati-Khoshghalb et al. 2019] Salavati-Khoshghalb, Majid ; Gendreau, Michel ; Jabali, Ola ; Rei, Walter: An exact algorithm to solve the vehicle routing problem with stochastic demands under an optimal restocking policy. In: *European Journal of Operational Research* 273 (2019), Nr. 1, p. 175–189.

[Sampaio et al. 2019] Sampaio, Afonso ; Kinable, Joris ; Veelenturf, Lucas P. ; Van Woensel, Tom: A scenario-based approach for the vehicle routing problem with roaming delivery locations under stochastic travel times. In: *Optimization Online* (2019).

[Sluijk et al. 2022] Sluijk, Natasja ; Florio, Alexandre M. ; Kinable, Joris ; Dellaert, Nico ; Van Woensel, Tom: Two-echelon vehicle routing problems: A literature review. In: *European Journal of Operational Research* (2022).

[Soeffker et al. 2021] Soeffker, Ninja ; Ulmer, Marlin W. ; Mattfeld, Dirk C.: Stochastic dynamic vehicle routing in the light of prescriptive analytics: A review. In: *European Journal of Operational Research* (2021).

[Song et al. 2020] Song, Yongjia ; Ulmer, Marlin W. ; Thomas, Barrett W. ; Wallace, Stein W.: Building trust in home services—stochastic team-orienteering with consistency constraints. In: *Transportation Science* 54 (2020), Nr. 3, p. 823–838.

[Spliet et al. 2018] Spliet, Remy ; Dabia, Said ; Van Woensel, Tom: The time window assignment vehicle routing problem with time-dependent travel times. In: *Transportation Science* 52 (2018), Nr. 2, p. 261–276.

[Spliet and Desaulniers 2015] Spliet, Remy ; Desaulniers, Guy: The discrete time window assignment vehicle routing problem. In: *European Journal of Operational Research* 244 (2015), Nr. 2, p. 379–391.

[Spliet and Gabor 2015] Spliet, Remy ; Gabor, Adriana F.: The time window assignment vehicle routing problem. In: *Transportation Science* 49 (2015), Nr. 4, p. 721–731.

[Statista 2020] Statista: *eCommerce - Weltweit*. 2020. – URL `https://de.statista.com/outlook/dmo/ecommerce/weltweit`. – Online; accessed 5 May 2022.

[Subramanyam et al. 2018] Subramanyam, Anirudh ; Wang, Akang ; Gounaris, Chrysanthos E.: A scenario decomposition algorithm for strategic time window assignment vehicle routing problems. In: *Transportation Research Part B: Methodological* 117 (2018), p. 296–317.

[Sungur et al. 2010] Sungur, Ilgaz ; Ren, Yingtao ; Ordóñez, Fernando ; Dessouky, Maged ; Zhong, Hongsheng: A model and algorithm for the courier delivery problem with uncertainty. In: *Transportation science* 44 (2010), Nr. 2, p. 193–205.

[Ulmer and Streng 2019] Ulmer, Marlin W. ; Streng, Sebastian: Same-day delivery with pickup stations and autonomous vehicles. In: *Computers & Operations Research* 108 (2019), p. 1–19.

[velove 2022] velove: `https://www.velove.se/`. 2022. – Online; accessed 05 May 2022.

[Visser and Savelsbergh 2019] Visser, Thomas ; Savelsbergh, Martin: Strategic time slot management: A priori routing for online grocery retailing. 2019. – Research Report.

[Zhen et al. 2020] ZHEN, Lu ; LV, Wenya ; WANG, Kai ; MA, Chengle ; XU, Ziheng: Consistent vehicle routing problem with simultaneous distribution and collection. In: *Journal of the Operational Research Society* 71 (2020), Nr. 5, p. 813–830.

A Appendix

A.1 Clustered Demand Pattern

In this section we explain in more detail the clustered demand pattern. As shown in Figure 4, the city is divided into city centre, industrial area, and residential areas. Pickup and delivery customer locations are sampled as follows. Half of the pickup customer locations are generated uniformly over the city centre, the other half is generated uniformly over the industrial zone. This is motivated by shops and companies being located in these areas usually. As companies also might order goods or people might order parcels to their work place instead of their home, a fourth of the delivery customer locations is sampled uniformly within the industrial zone. Still, the majority of parcels is ordered to private households. Therefore the remaining four quarters of delivery customer locations are distributed uniformly in the residential areas.

A.2 Detailed Analysis of Different Consensus Functions

In this section, we investigate the different consensus functions in ore detail. We are interested in which consensus function performs best for each service design and demand pattern. For this purpose, Figure 10 shows the average relative gap to the daily solution of the MSA with different consensus functions as well as the fixed solution per service design and demand pattern.

For *uniform* instances with instant delivery, `reachability` yields the best results. On average, it has a gap of 16.18% to the daily solution. Also for the *clustered* demand pattern `reachability` performs best for instant delivery, showing an average gap of 10.58% to the daily solution. If time is short as in the instant delivery design, it is both important to drive efficient routes (i.e. to not loose too much time on the streets) and to drive along arcs connecting relevant micro-hubs. Due to the limited service time horizon, there is not a path from every micro-hub to any other micro-hub – `reachability` thus offers a good criterion to find a solution that shows high similarity in the reachability between micro-hubs. When investigating the different consensus functions for instant delivery, we further find that the consistent tour produced by `reachability` has less stops than the ones produced by other consensus functions or the fixed solution. Since flexibility with respect to time is limited, fewer stops – and possibly waiting at micro-hubs – features the spacial consolidation such that several parcels can be loaded during one stop. Further, in the master tours produced by `reachability`, the vehicle waits up to three times at a micro-hub. Using other consensus functions, the vehicle waits either never or at most once per route.

For fast and same-day delivery reachability is not a good comparison criterion any more since at time 0 every micro-hub is connected to every other micro-hub. For *uniform* instances with fast delivery, `direct` delivers the best average objective values, deviating by 6.98% from the daily solution on average. On these instances, the fixed solution performs particularly bad, leading to average objective values about 36.24% worse than in the daily solution. On *clustered* demand, consensus function `time` shows the smallest deviation of 8.21% from the daily solution with `direct` and `order` following closely. Since in the fast delivery design, service time horizon as well as delivery time promise for parcel requests are longer, time is less restrictive in the delivery process as in instant delivery. Up to 79.66% of all parcel requests can be fulfilled on these instances. Analysing the individual solutions on *uniform* instances with fast delivery, we find that many show a circular route pattern (see Section 5.3.1). Since `direct` compares direct neighbours within routes, this consensus function is likely to detect such a circular structure well. In the *clustered* demand pattern, some micro-hubs show to have more pickup or delivery requests, respectively. It hence makes a difference when and in which order micro-hubs are visited. This explains the good performance of `time`, `direct` and `order` on these instances. While `time` controls at which hour a micro-hub is visited, `direct` compares direct neighbourhoods, and `order` considers the full sequence of micro-hubs in a tour. Since those three consensus functions yield similar average objective

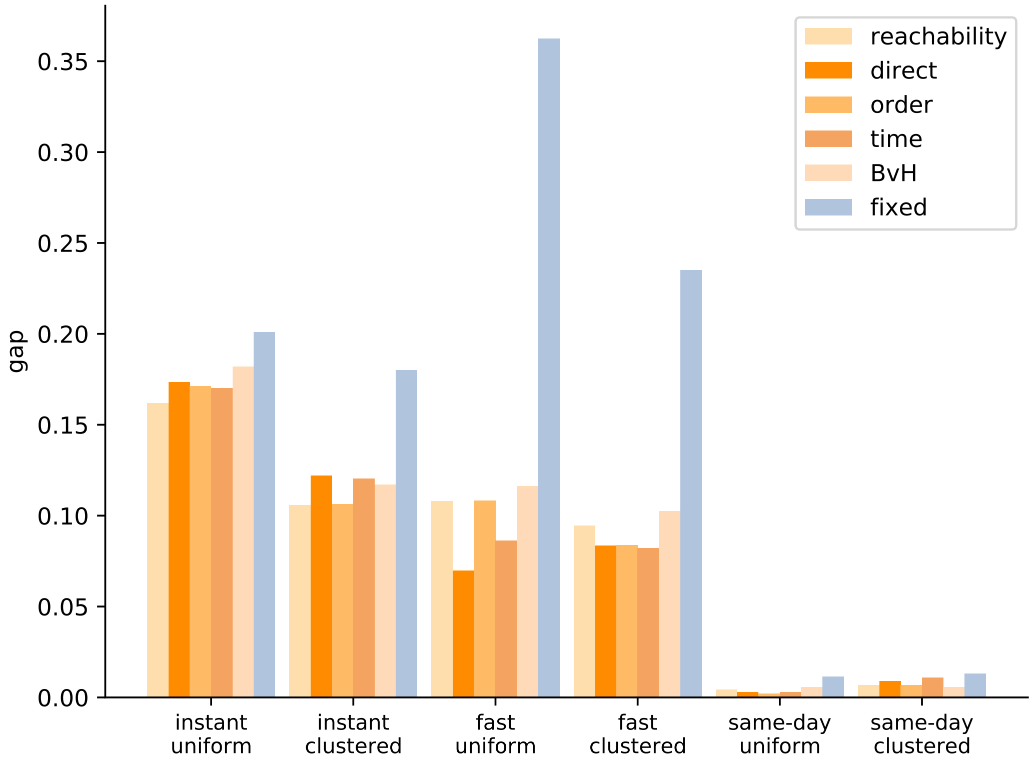


Figure 10: Average relative gap to daily solution per consensus function and instance size.

values, this indicates that all three characteristics are important to find a robust consistent tour for fast delivery on clustered demand.

For same-day delivery, `order` performs best for *uniform*, and `BvH` for *clustered* demand, yielding solutions that deviate by 0.20%, and 0.58% from the daily solution on average. From Figure 10 we see that MSA performs much better on these large delivery time promises than on the smaller ones. This can be explained by the fact that the more time is available, the easier it becomes to transport the same amount of parcels. For *uniform* demand, the solutions produced by `order` are at most 1.26 parcels behind the daily solution; on average it differs by only 0.11 parcels. On average, the consistent tour produced by MSA deviates by just 0.35% from a non-consistent scenario-dependent route. Further, on instances with *uniform* demand and 80 parcels, we are able to fulfil more than 95.48% of all parcel requests with any consensus function. For *clustered* demand we obtain results that are even closer to the daily solution: on average of all consensus functions, the solutions produced by the MSA differ by only 0.51 parcels from the daily solution on average. In this case, the MSA thus creates a consistent master tour that reduces solution quality by only 0.46% on average compared to a daily re-optimised inconsistent routing strategy.

Otto von Guericke University Magdeburg
Faculty of Economics and Management
P.O. Box 4120 | 39016 Magdeburg | Germany

Tel.: +49 (0) 3 91 / 67-1 85 84
Fax: +49 (0) 3 91 / 67-1 21 20

www.fww.ovgu.de/femm

ISSN 1615-4274