



AARHUS UNIVERSITY



Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the accepted manuscript version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

Kiilerich, L., & Wøhlk, S. (2018). New large-scale data instances for CARP and new variations of CARP. *INFOR: Information Systems and Operational Research*, 56(1), 1-32. <https://doi.org/10.1080/03155986.2017.1303960>

Publication metadata

Title: New large-scale data instances for CARP and new variations of CARP
Author(s): Kiilerich, L., & Wøhlk, S.
Journal: *INFOR: Information Systems and Operational Research*
DOI/Link: [10.1080/03155986.2017.1303960](https://doi.org/10.1080/03155986.2017.1303960)
Document version: Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

If the document is published under a Creative Commons license, this applies instead of the general rights.

New Large-Scale Data Instances for CARP and New Variations of CARP

Lone Kiilerich^a and Sanne Wøhlk^a

^aCluster for Operations Research And Logistics, Department of Economics and Business Economics, Aarhus University, Denmark

ARTICLE HISTORY

Compiled July 24, 2017

ABSTRACT

The Capacitated Arc Routing Problem (CARP) captures important aspects of real-life problems and has been studied extensively over the past two decades. Based on a waste collection project, we introduce a number of new CARP variations. We first present three multi-compartment CARP variations of different levels of complexity regarding compartments and where one incorporates a time horizon. We then present a variation that seeks to coordinate vehicles over a planning horizon such that the vehicles that collect different waste fractions from the same households do so on the same day of the week. Finally, the semi-periodic CARP takes into account that the households on a street, providing the demand of the edge, may not request waste collection at the same interval. We present large-scale instances both for the classical CARP and for the five new problems. The instances are based on real-life networks and waste data from five areas in Denmark and cover rural as well as urban areas. The largest instances contain more than 10 thousand nodes. We give detailed information about the construction of the instances from the real-life data, and explain how they can be used to perform scenario analyses.

KEYWORDS

Capacitated arc routing problem; Waste collection; Real-life benchmark instances; New variations.

1. Introduction

The motivation for writing this paper is twofold: Firstly, through collaboration with a number of waste responsible units, we have become aware of a number of CARP variations which are highly relevant in practice, but have not been studied in the academic literature. We therefore take this opportunity to present these problems, their relevance, and how they can be used for comparison of different waste collection strategies. Secondly, it is well known in the arc routing community that the classical benchmark instances, the largest being the instances presented in Brandão and Eglese (2008) with 255 nodes and 375 edges, are significantly smaller than most real-life instances of arc routing. We therefore present new large-scale real-life data (with thousands of nodes) both for the CARP and for the new variations, and make this data available for the research community as benchmark data.



Figure 1.: The counties underlying the data.

This paper is part of a larger project on transportation problems within waste collection. In the project, we collaborate with a number of counties and partially state-owned companies responsible for waste collection. They provide us with knowledge of the problems they encounter and with valuable data. Figure 1 provides a graphical overview of our collaborators as regards curbside collection. Starting from the top left, Reno Djurs I/S (the counties of North Djurs (N) and South Djurs (S)) (Reno Djurs 2017) and Renosyd I/S (the counties of Skanderborg and Odder (K)) (Renosyd 2017) are relatively rural areas with towns and villages of various sizes. Odense Renovation A/S (Odense county (O)) (Odense Renovation A/S 2017) operates in the third largest city in Denmark that contains an urban mixture of single-family houses and apartment block areas. Finally, Frederiksberg county (F) (Frederiksberg county 2017) is an independent county covering part of central Copenhagen and has a more traditional city-like structure. Letters in brackets refer to the naming of our data, which is described in Section 4.

In this paper, we present a total of five new CARP variations that can be grouped into three types. These are represented by the columns in Figure 2, and a shaded box indicates a problem which includes a time perspective. The lines in the figure represent extensions of the problems, where the simplest problem is the classical CARP at the top. All five variations are based on real-life situations encountered by our collaborators. Even though the areas do include some one-way streets, we have agreed that an undirected version of the problems will suffice.

The first type of problems are multi-compartment problems. With the increasing level of waste sorting, multiple waste fractions must be collected from each household. The first variation of multi-compartment problems arises from the situation where each household has a single bin with a number of compartments for sorted waste fractions. The bin is emptied by a multi-compartment vehicle that has a number of compartments matching those of the bin. Practical applications within waste collection contain up to four compartments. The same problem also arises when separate bins are collected by a single vehicle with the compartments pre-assigned to waste fractions. We refer to this problem as No-Split Multi-Compartment CARP (No-Split MC-CARP) and it is studied in Section 3.1.1.

When the number of waste fractions to be collected from each household exceeds the number of compartments in the vehicle or when additional flexibility is wanted, the

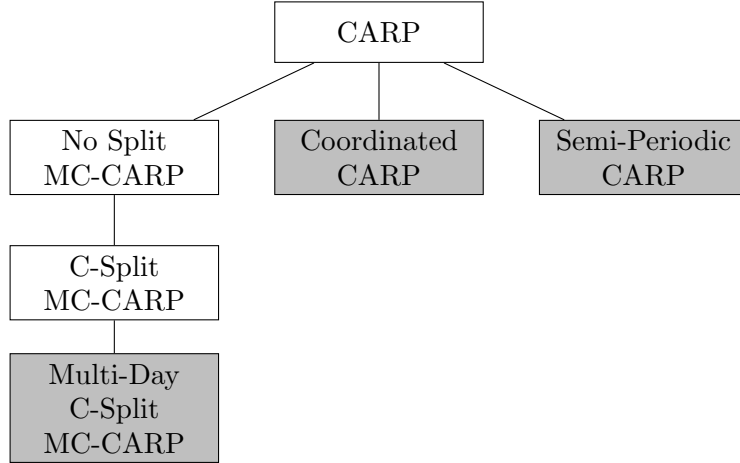


Figure 2.: Overview of the problems.

Commodity-Split Multi-Compartment CARP (C-Split MC-CARP) is more relevant. This problem is presented in Section 3.1.2 and involves, besides the routing aspect, the choice of waste fraction to be collected in each compartment of each vehicle. This also means that each household may be serviced by multiple vehicles.

Collection of waste is, in fact, a multi-day problem. While this is but an unimportant issue when each household is serviced by a single vehicle and compartments are pre-assigned to waste fractions (because the set of routes can be distributed over the days), it does become important when multiple vehicles service each household and the compartment assignment is part of the decision process. To avoid having to thoroughly clean every compartment at the end of each day, the same compartment should be used for collection of the same waste fraction every day in practice. The Multi-Day Commodity-Split Multi-Compartment CARP (Multi-Day C-Split MC-CARP) presented in Section 3.1.3 takes this timing perspective into consideration.

The time perspective is also included in the next type of problem, the coordinated CARP, which is considered in Section 3.2. When multiple vehicles collect different fractions of waste from each household, the services should be coordinated to take place on the same set of days. In the Coordinated CARP (C-CARP), this is considered in a setting with single compartment vehicles.

The last type of problem, the Semi-Periodic CARP, is relevant in cases where the households on a street segment may not request the same service frequency, even for the same waste fraction. For instance, some of our collaborators have customers who require collection of general waste twice a week, once a week, or every second week. This results in a periodic-like problem, which we call Semi-Periodic CARP (SP-CARP) which is studied in Section 3.3.

When the fleet of vehicles is constructed, cost of the individual vehicles could be considered, these include investment costs and maintenance. This implies that the fleet size and choice of vehicle have a direct impact on the cost of servicing demand. Problems that include a time horizon should therefore preferably have the routes spread evenly on all days in the planning horizon and not clustered on a few days. As the cost of a vehicle is often large compared to the routing costs, it might even be preferred to create the routes in such a way that the cost of needed vehicles is minimized. We therefore formulate a bi-criteria problem in cases where a planning horizon is included. This bi-criteria problem first seeks to minimize the cost of vehicles

used and secondarily the routing costs.

Compression is an important aspect of waste collection and we therefore discuss this issue before presenting the models. Compression relates to the fact that 100 liters of waste in a waste bin do not require 100 liters of space in the waste truck. If we take a mixed waste fraction that contains glass, paper and organic waste to be delivered to an incineration plant, the waste can actually be compressed by a factor of 5 to 6 and therefore only requires 17-20 liters in the vehicle. The compression comes from a combination of load of the other waste in the vehicle and mechanical compression. On the other hand, a waste fraction containing a mix of glass, metal, and plastic to be brought to a sorting facility is not being compressed beyond natural compression because it complicates the subsequent sorting. From a modeling perspective, these differences in compression factors for different waste fractions mean, in effect, that the actual compartment capacity depends on the waste fraction it is assigned to. Furthermore, different vehicles has different compression abilities. The compression factor, and thereby the effective compartment size, will therefore depend on both the waste fraction and the vehicle. This is reflected in all models in Section 3.

The main contributions of this paper are:

- (1) We introduce five new CARP variations, all of which are encountered in real-life waste collection.
- (2) We provide new large-scale instances for these five variations as well as for the CARP. The instances are based on real-life waste data from five areas in Denmark.
- (3) We describe how the new instances can be used for comparison and scenario analyses.

The remainder of this paper is organized as follows. In Section 2, we review the related literature and in Section 3, we describe the five new CARP variations. Section 4 is devoted to a description of the new instances and how they are generated, and Section 5 explains how scenario analysis can be performed across the instances. Finally, Section 6 offers some concluding remarks, and the appendices give detailed information about the instances.

2. Related Literature

The CARP was introduced by Golden and Wong (1981) and has a wide range of applications. The one in focus in this article is urban waste collection. As different countries and cities handle waste collection in many different ways, the original CARP has been extended to several more complicated versions that seek to model the practice of waste handling. This section will give a brief introduction to these extensions. For a thorough review of studied arc routing problems until 2014 in this context see Corberán and Laporte (2015, chapter 19).

In many real-world applications it is not possible to traverse a street in both directions or to collect waste from both sides of the road when traversing in one direction. This has led to the use of Directed CARP (Amponsah and Salhi 2004; McBride 1982; Mourão and Almeida 2000), where the graph contains arcs rather than edges. A version where the network contains arcs and edges that can be completely serviced when traversing in one direction (zigzag service) is denoted Mixed CARP (Bautista et al. 2008; Belenguer et al. 2006; Constantino et al. 2015; Coutinho-Rodrigues et al. 1993; Ghiani et al. 2005; Gouveia et al. 2010; Mourão and Amado 2005; Mourão et al. 2009;

Santos et al. 2011). A third version that relates to specific characteristics of the road network is denoted windy CARP, here the cost of traversing an edge depends on the direction in which it is traversed. This can be used to model hills and wind direction, for instance (Corberán et al. 2011). A further complication in real-world waste collection is forbidden or unwanted turns such as U-turns or left turns (Bautista et al. 2008; Coutinho-Rodrigues et al. 1993; McBride 1982; Mourão et al. 2009; Santos et al. 2011).

In the real-world, the dump site and the vehicle depot might not be at the same location or there may be several dump sites of which the depot is one. This presents a CARP with intermediate facilities in which the vehicles can be unloaded at any of these nodes (Amponsah and Salhi 2004; Ghiani et al. 2005, 2010; Polacek et al. 2008; Santos et al. 2011).

The characteristics of the allowed routes may differ this lead to the following extensions. Sometimes vehicles can be filled several times. In this case a route consists of more than one visit to a dump site. This is often the case if there are several dump sites, but it is also relevant in cases where demand is relatively large compared to capacity of the vehicle (Amponsah and Salhi 2004; Ghiani et al. 2001; Mourão and Almeida 2000; Mourão et al. 2009; Polacek et al. 2008; Santos et al. 2011). The wish to be able to make the workers who are assigned to a route responsible for the quality of their work is another issue that influences route planning. Therefore, non-overlapping routes may be desirable. This implies that the serviced edges of a route should be somehow connected and the graph has to be divided up into areas that should be serviced by one vehicle (Constantino et al. 2015). Another desired constraint on the routes is that they are somehow balanced. When working with districts this is obtained by making the districts balanced (Constantino et al. 2015). A less restrictive way of ensuring some balance among routes is to constrain the allowed length or time duration of the routes. This could also be relevant if the routes are to be serviced within one work shift (Ghiani et al. 2005, 2010; Mourão et al. 2009; Polacek et al. 2008; Santos et al. 2011).

The vehicles available for waste collection are not necessarily identical. This gives rise to a CARP version with heterogeneous vehicles (Ghiani et al. 2005; Del Pia and Filippi 2006). When vehicles differ, it might be the case that some streets cannot be serviced by all vehicle types, either because the vehicle cannot traverse the given edge or due to technical reasons that prevent bins on the edge to be emptied by some vehicle types. These cases are also investigated in Ghiani et al. (2005); Del Pia and Filippi (2006). In Del Pia and Filippi (2006) the small vehicles that can service narrow edges denoted satellite vehicles, are emptied into large vehicles. This implies that both vehicles have to be at the same location for the period of time it takes to empty the small vehicle into the large one. Another complication regarding service is when some edges can only be serviced during a given time window, for instance in order to avoid traffic congestion on busy streets during rush hour (Ghiani et al. 2005).

Two versions modify the objective function in order to take real-world complications into consideration. The first contains a punishment for uncollected waste (Amponsah and Salhi 2004) and arises when the available vehicles cannot service all demand and uncollected waste implies both health risk and annoyance for the citizens affected. The second version is when the waste collection is maintained by a private company that does not have to service all demand. Then the company will be more interested in profit than cost. We can model this using a prize collecting arc routing problem with the constraint that if one edge is serviced, then all edges in its connected demand component should be serviced by the same vehicle (Aráoz et al. 2013; Corberán et al.

2011).

The last extension takes into account that waste is not collected every day and that the frequency between service is not necessarily the same on all edges. This gives rise to Periodic CARP (PCARP), where demand occurs every day and is accumulated until service of the edge. This version is studied in a waste collection setting in Mei et al. (2011), Dos Santos et al. (2016), and Zhang et al. (2016).

3. New CARP Variations

All models are based on the first model for the CARP, presented in Golden and Wong (1981). However, we use slightly different variable definitions. We state the model for the CARP below for future reference and for clarification.

The CARP is defined on an undirected connected graph $G = (\mathcal{N}, \mathcal{E})$, where \mathcal{N} is the set of nodes and \mathcal{E} is the set of edges. The edges are defined as ordered pairs (i, j) with $i, j \in \mathcal{N}$ and $i < j$. The graph is assumed not to contain loops or multiple edges between the same pair of nodes. With every edge, $(i, j) \in \mathcal{E}$ is associated a traversal cost $c_{ij} > 0$. A subset of the edges, $\mathcal{E}_R \subseteq \mathcal{E}$ are required. For these edges, $q_{ij} > 0$ denotes the demand of edge $(i, j) \in \mathcal{E}_R$. A special depot node, $dep \in \mathcal{N}$, holds a homogeneous set of vehicles \mathcal{K} . Each vehicle has a capacity of W and a compression factor of γ . To model the CARP, we define two types of variables. For every required edge $(i, j) \in \mathcal{E}_R$, we define a binary service variable y_{ij}^k as follows:

$$y_{ij}^k = \begin{cases} 1 & \text{if edge } (i, j) \in \mathcal{E}_R \text{ is serviced by vehicle } k \in \mathcal{K}, \\ 0 & \text{otherwise.} \end{cases}$$

For every edge $(i, j) \in \mathcal{E}$, we define two binary traversal variables x_{ij}^k and x_{ji}^k as follows:

$$x_{ij}^k = \begin{cases} 1 & \text{if edge } (i, j) \in \mathcal{E} \text{ is traversed from } i \text{ to } j \text{ by vehicle } k \in \mathcal{K}, \\ 0 & \text{otherwise} \end{cases}$$

and

$$x_{ji}^k = \begin{cases} 1 & \text{if edge } (i, j) \in \mathcal{E} \text{ is traversed from } j \text{ to } i \text{ by vehicle } k \in \mathcal{K}, \\ 0 & \text{otherwise.} \end{cases}$$

With this, the CARP is described as

$$\min \sum_{k \in \mathcal{K}} \sum_{(i,j) \in \mathcal{E}} c_{ij} (x_{ij}^k + x_{ji}^k)$$

$$\text{st. } \sum_{k \in \mathcal{K}} y_{ij}^k = 1 \quad \forall (i,j) \in \mathcal{E}_R \quad (1)$$

$$\sum_{(i,j) \in \mathcal{E}} q_{ij} y_{ij}^k \leq \gamma W \quad \forall k \in \mathcal{K} \quad (2)$$

$$x_{ij}^k + x_{ji}^k \geq y_{ij}^k \quad \forall k \in \mathcal{K}, (i,j) \in \mathcal{E}_R \quad (3)$$

$$\sum_{j \in \mathcal{N}: (i,j) \in \mathcal{E}} (x_{ij}^k - x_{ji}^k) = 0 \quad \forall k \in \mathcal{K}, i \in \mathcal{N} \quad (4)$$

$$\text{Subtour elimination constraints} \quad (5)$$

$$x_{ij}^k, x_{ji}^k \in \{0, 1\} \quad \forall k \in \mathcal{K}, (i,j) \in \mathcal{E} \quad (6)$$

$$y_{ij}^k \in \{0, 1\} \quad \forall k \in \mathcal{K}, (i,j) \in \mathcal{E}_R. \quad (7)$$

The objective function states that we seek to minimize the total traversal cost. Constraints (1) ensure that all required edges are serviced by exactly one vehicle and (2) are the vehicle capacity constraints. Constraints (3) are coupling constraints forcing a vehicle to traverse any edge that it services. Jointly, the route continuity constraints (4), the subtour elimination constraints (5) ensure that every route starts and ends in the depot node, and the domain constraints for the traversal variables (6) comprise the route topology constraints, and finally, (7) define the domain of the service variables.

3.1. Multi-Commodity CARP

In this section, we describe the three versions of MC-CARP in more detail. As the notation is quite extensive, it is summarized in Table 1. Throughout the section, we consider a set of waste fractions \mathcal{F} . These will also be denoted commodities. For each required edge $(i,j) \in \mathcal{E}_R$, we denote by q_{ij}^f the demand of the edge for waste fraction $f \in \mathcal{F}$. Note that a required edge need not request service for all waste fractions, i.e. $q_{ij}^f \geq 0, \forall f \in \mathcal{F}, \forall (i,j) \in \mathcal{E}_R$, but each required edge must request service for at least one of the waste fractions such that we have $\sum_{f \in \mathcal{F}} q_{ij}^f > 0, \forall (i,j) \in \mathcal{E}_R$. We use $\mathcal{E}_R^f \subseteq \mathcal{E}_R$ to denote the set of required edges with respect to waste fraction $f \in \mathcal{F}$, i.e. $\mathcal{E}_R^f = \{(i,j) \in \mathcal{E}_R : q_{ij}^f > 0\}$. With this, we can formulate the three versions of the MC-CARP.

3.1.1. No-Split Multi-Commodity CARP

In the No-Split version of MC-CARP, all waste fractions from an edge must be collected by a single vehicle, or equivalently, all commodities demanded by an edge must be serviced by a single vehicle. In this problem, the vehicles are identical, each having $|\mathcal{F}|$ compartments. We define W^f to be the capacity of the vehicle compartment for waste fraction $f \in \mathcal{F}$ and denote the compression factor for fraction $f \in \mathcal{F}$ by γ^f . Because it is not allowed to split the service of a single edge between several vehicles, we can use the same service variables as for the CARP.

Index Sets and Indices

\mathcal{N}	The set of nodes, indexed by i and j .
\mathcal{E}	The set of edges, indexed by the ordered pair (i, j) with $i < j$.
\mathcal{E}_R	The set of required edges, i.e. $\mathcal{E}_R = \{(i, j) \in \mathcal{E} : q_{ij} > 0\}$.
\mathcal{E}_R^f	The set of required edges wrt. $f \in \mathcal{F}$, i.e. $\mathcal{E}_R = \{(i, j) \in \mathcal{E} : q_{ij}^f > 0\}$.
\mathcal{K}	The set of vehicles, indexed by k .
\mathcal{M}_k	The set of compartments in vehicle $k \in \mathcal{K}$.
\mathcal{F}	The set of waste fractions, indexed by f .
\mathcal{T}	The set of days in the time horizon, indexed by t .

Parameters

$W; W^f$	Capacity of vehicle; Capacity of vehicle compartment for waste fraction $f \in \mathcal{F}$.
W^{km}	Capacity of compartment $m \in \mathcal{M}_k$ of vehicle $k \in \mathcal{K}$.
C^k	Cost of including vehicle $k \in \mathcal{K}$ in the fleet.
c_{ij}	Traversal cost of edge $(i, j) \in \mathcal{E}$, $c_{ij} > 0$.
$q_{ij}; q_{ij}^f$	Demand of edge $(i, j) \in \mathcal{E}$, $q_{ij} \geq 0$; Demand for waste fraction $f \in \mathcal{F}$ of edge $(i, j) \in \mathcal{E}$.
γ^f	Compression factor for waste fraction $f \in \mathcal{F}$.
γ^{fk}	Compression factor for waste fraction $f \in \mathcal{F}$ when collected by vehicle $k \in \mathcal{K}$.

Binary Variables

v^k	1 if vehicle $k \in \mathcal{K}$ is used in solution, and zero otherwise.
x_{ij}^k	1 if vehicle $k \in \mathcal{K}$ traverses edge $(i, j) \in \mathcal{E}$ in direction from i to j , and zero otherwise. x_{ji}^k is 1 if vehicle $k \in \mathcal{K}$ traverses the same edge (i, j) in the direction from j to i , and zero otherwise.
x_{ij}^{kt}	Extension of x_{ij}^k to every day $t \in \mathcal{T}$.
y_{ij}^k	1 if vehicle $k \in \mathcal{K}$ services edge $(i, j) \in \mathcal{E}_R$, and zero otherwise.
y_{ij}^{fk}	1 if $k \in \mathcal{K}$ services waste fraction $f \in \mathcal{F}$ of $(i, j) \in \mathcal{E}_R$, and zero otherwise.
y_{ij}^{fkt}	Extension of y_{ij}^{fk} to every day $t \in \mathcal{T}$.
z_{ij}^{fkm}	1 if vehicle $k \in \mathcal{K}$ uses compartment $m \in \mathcal{M}_k$ for waste fraction $f \in \mathcal{F}$.

Table 1.: Overview of notation for the three MC-CARP variations.

The No-Split MC-CARP can then be modeled as follows.

$$\begin{aligned} \min \quad & \sum_{k \in \mathcal{K}} \sum_{(i,j) \in \mathcal{E}} c_{ij} (x_{ij}^k + x_{ji}^k) \\ \text{st.} \quad & \sum_{(i,j) \in \mathcal{E}} q_{ij}^f y_{ij}^k \leq \gamma^f W^f \quad \forall f \in \mathcal{F}, k \in \mathcal{K} \end{aligned} \quad (8)$$

$$\text{Constraints (1), (3) – (7).} \quad (9)$$

Compared to the CARP, the only difference in this model is that for the No-Split MC-CARP, we must ensure that the vehicle capacity is respected for every compartment of the vehicle.

3.1.2. C-Split Multi-Commodity CARP

The C-Split version of the problem is significantly more flexible than the No-Split.

In this problem, we allow each edge to be serviced by more than one vehicle. In fact, it may not be possible to service all demand of an edge by a single vehicle. In some areas, 6 or more waste fractions must be collected from a single edge even though the most advanced waste collection vehicle that we know of only collects 4 separated waste fractions. This adds an extra dimension of complexity as it must now be determined which waste fractions should be collected by which vehicle and in which compartment, besides determining which vehicle should service each edge and which subset of commodities of the edge. We still require the full amount of each waste fraction of the edge to be collected by a single vehicle.

The vehicle fleet \mathcal{K} is heterogeneous and for each vehicle $k \in \mathcal{K}$ we define \mathcal{M}_k to be the set of compartments. The capacity of compartment $m \in \mathcal{M}_k$ in vehicle $k \in \mathcal{K}$ is given by W^{km} and vehicle $k \in \mathcal{K}$ can compress fraction $f \in \mathcal{F}$ by a factor γ^f .

To relate the compartments of the vehicles to the waste fractions, we define a compartment variable for each waste fraction and compartment combination as follows:

$$z^{fkm} = \begin{cases} 1 & \text{if vehicle } k \in \mathcal{K} \text{ collects waste fraction } f \in \mathcal{F} \text{ in compartment } m \in \mathcal{M}_k, \\ 0 & \text{otherwise.} \end{cases}$$

The service variables are redefined to include information on the waste fraction and are defined $\forall f \in \mathcal{F}$, $(i, j) \in \mathcal{E}_R^f$, and $\forall k \in \mathcal{K}$.

$$y_{ij}^{fk} = \begin{cases} 1 & \text{if waste fraction } f \in \mathcal{F} \text{ on edge } (i, j) \in \mathcal{E}_R^f \text{ is serviced by vehicle } k \in \mathcal{K}, \\ 0 & \text{otherwise.} \end{cases}$$

Then the model becomes

$$\min \sum_{k \in \mathcal{K}} \sum_{(i,j) \in \mathcal{E}} c_{ij}(x_{ij}^k + x_{ji}^k)$$

$$\text{st. } \sum_{k \in \mathcal{K}} y_{ij}^{fk} = 1 \quad \forall f \in \mathcal{F}, (i,j) \in \mathcal{E}_R^f \quad (10)$$

$$\sum_{(i,j) \in \mathcal{E}} q_{ij}^f y_{ij}^{fk} \leq \sum_{m \in \mathcal{M}_k} z^{fkm} \gamma^{fk} W^{km} \quad \forall f \in \mathcal{F}, k \in \mathcal{K} \quad (11)$$

$$\sum_{f \in \mathcal{F}} z^{fkm} \leq 1 \quad \forall k \in \mathcal{K}, m \in \mathcal{M}_k \quad (12)$$

$$|\mathcal{M}_k|(x_{ij}^k + x_{ji}^k) \geq \sum_{f \in \mathcal{F}} y_{ij}^{fk} \quad \forall k \in \mathcal{K}, (i,j) \in \mathcal{E}_R \quad (13)$$

$$\text{Constraints (4) - (6)} \quad (14)$$

$$y_{ij}^{fk} \in \{0, 1\} \quad \forall f \in \mathcal{F}, k \in \mathcal{K}, (i,j) \in \mathcal{E}_R^f \quad (15)$$

$$z^{fkm} \in \{0, 1\} \quad \forall f \in \mathcal{F}, k \in \mathcal{K}, m \in \mathcal{M}_k. \quad (16)$$

Constraints (10) ensure that all waste fractions are collected from every edge. Compared to the similar constraint (1) for the No-Split MC-CARP, it is now necessary to ensure service for each commodity by separate constraints. The vehicle capacities are respected by constraints (11). Here, the left hand side denotes the total amount of waste fraction f collected by vehicle k and the right hand side is the total capacity assigned to fraction f in the vehicle. If several compartments of a vehicle are used for the same waste fraction, the right hand side provides their joint capacity. If waste fraction f is not collected by vehicle k , the right hand side becomes zero, not allowing the vehicle to collect f . Constraints (12) ensure that each compartment of each vehicle is only used for one waste fraction. Constraints (13) are the counterparts of (3), ensuring that a vehicle traverses an edge if it services any of the demand of that edge. Finally, the route topology constraints (4)-(6) are unchanged from the CARP, and (15) and (16) define the domains of the service and compartment variables.

3.1.3. Multi-Day Commodity-Split MC-CARP

With this third MC-CARP, we move closer to the problem observed in the real-world and further away from the classical CARP. When we look at a daily problem involving multiple waste fractions, we are faced with either a No-Split MC-CARP or a C-Split MC-CARP, but over time, this is not necessarily the case. In fact, if a given compartment of a certain vehicle is used to collect one waste fraction today, that vehicle should then collect the same waste fraction in that compartment tomorrow and the next day and so on. So, when we set free the assignment of waste fractions to compartments as done through the z^{fkm} variables for the C-Split MC-CARP, the decision needs to be permanent over the planning horizon. This is what we consider in the model in this section. Alternatively, a two phase strategy can be applied; first assigning waste fractions to vehicles and compartments at a tactical level and secondly solving a number of independent No-Split MC-CARP at the operational level.

To include the time perspective, we define the time horizon as a set of days \mathcal{T} , indexed by t . Each edge must be serviced once during the time horizon and q_{ij}^f denotes the demand for waste fraction f . In Sections 3.2 and 3.3, we consider problems where

the time perspective results in multiple services of each edge during the time horizon. Let C^k , $k \in \mathcal{K}$ be the cost of including vehicle k in the fleet.

To model the Multi-Day Commodity-Split MC-CARP, we add a time index to the traversal and the service variables. As the assignment of waste fractions to compartments is fixed over the time horizon, the compartment variables are unchanged. For $\forall (i, j) \in \mathcal{E}$, $k \in \mathcal{K}$, and $\forall t \in \mathcal{T}$, we have

$$x_{ij}^{kt} = \begin{cases} 1 & \text{if edge } (i, j) \in \mathcal{E} \text{ is traversed from } i \text{ to } j \text{ by vehicle } k \in \mathcal{K} \text{ at day } t, \\ 0 & \text{otherwise,} \end{cases}$$

$$x_{ji}^{kt} = \begin{cases} 1 & \text{if edge } (i, j) \in \mathcal{E} \text{ is traversed from } j \text{ to } i \text{ by vehicle } k \in \mathcal{K} \text{ at day } t, \\ 0 & \text{otherwise,} \end{cases}$$

and $\forall f \in \mathcal{F}$, $(i, j) \in \mathcal{E}_R^f$, $k \in \mathcal{K}$, and $\forall t \in \mathcal{T}$, we have

$$y_{ij}^{fkt} = \begin{cases} 1 & \text{if fraction } f \in \mathcal{F} \text{ on } (i, j) \in \mathcal{E}_R^f \text{ is serviced by vehicle } k \in \mathcal{K} \text{ at day } t, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, to model the use of vehicles, we define the following for every vehicle $k \in \mathcal{K}$.

$$v^k = \begin{cases} 1 & \text{if vehicle } k \in \mathcal{K} \text{ is used on any route,} \\ 0 & \text{otherwise,} \end{cases}$$

The problem can then be described by the following model

$$\text{lex min} \left(\sum_{k \in \mathcal{K}} C^k v^k, \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \sum_{(i, j) \in \mathcal{E}} c_{ij} (x_{ij}^{kt} + x_{ji}^{kt}) \right)$$

$$\text{st. } \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} y_{ij}^{fkt} = 1 \quad \forall f \in \mathcal{F}, (i, j) \in \mathcal{E}_R^f \quad (17)$$

$$\sum_{(i,j) \in \mathcal{E}_R^f} q_{ij}^f y_{ij}^{fkt} \leq \sum_{m \in \mathcal{M}_k} z^{fkm} \gamma^{fk} W^{km} \quad \forall f \in \mathcal{F}, k \in \mathcal{K}, t \in \mathcal{T} \quad (18)$$

$$\sum_{f \in \mathcal{F}} z^{fkm} \leq v^k \quad \forall k \in \mathcal{K}, m \in \mathcal{M} \quad (19)$$

$$|\mathcal{M}_k| (x_{ij}^{kt} + x_{ji}^{kt}) \geq \sum_{f \in \mathcal{F}: q_{ij}^f > 0} y_{ij}^{fkt} \quad \forall k \in \mathcal{K}, t \in \mathcal{T}, (i, j) \in \mathcal{E}_R \quad (20)$$

$$\sum_{j \in \mathcal{N}: (i,j) \in \mathcal{E}} (x_{ji}^{kt} - x_{ij}^{kt}) = 0 \quad \forall k \in \mathcal{K}, t \in \mathcal{T}, i \in \mathcal{N} \quad (21)$$

$$\text{Subtour elimination constraints} \quad (22)$$

$$x_{ij}^{kt}, x_{ji}^{kt} \in \{0, 1\} \quad k \in \mathcal{K}, t \in \mathcal{T}, (i, j) \in \mathcal{E} \quad (23)$$

$$y_{ij}^{fkt} \in \{0, 1\} \quad \forall f \in \mathcal{F}, k \in \mathcal{K}, t \in \mathcal{T}, (i, j) \in \mathcal{E}_R^f \quad (24)$$

$$z^{fkm} \in \{0, 1\} \quad \forall f \in \mathcal{F}, k \in \mathcal{K}, m \in \mathcal{M}_k \quad (25)$$

$$v^k \in \{0, 1\} \quad \forall k \in \mathcal{K}, \quad (26)$$

where *lex min* refers to minimizing in lexicographical order, i.e. the objective functions are sequentially minimized in the order in which they are stated.

The objective function ensures that the number of vehicles is minimized at first and then secondly, the total travel distance is minimized. Constraints (17) ensure that every required waste fraction of every edge is serviced once during the time horizon. (18) extends (11) from the C-Split MC-CARP and ensures that the compartment capacities are respected every day, and (19) is an extension of (12) and assigns waste fractions to vehicle compartments for vehicles that are included in the solution. These constraints ensure that each compartment is used for the same fraction every day of the time horizon. Finally, (20)-(23) are direct extensions of the coupling constraints (3) and topology constraints (4)-(6) to every day, and (24) to (26) define the domains of the service, compartment, and used vehicle variables, respectively.

3.2. Coordinated CARP

In the previous section, we considered a number of problems related to the collection of multiple waste fractions. Given the increasing demand for waste sorting, each household may very likely need to have multiple waste bins. During our discussions with counties, waste collecting companies, and waste organizations in Denmark, we have learned that having up to four waste bins at a household, each with separate collection schedule, is a realistic scenario for the future. For a given citizen, this could mean that general and organic waste is collected every Monday, paper and cardboard every fourth Tuesday, glass and metal every third Thursday, and finally plastic every second Friday. This citizen thus faces the 12-week schedule seen in Figure 3 (in which weekends are removed).

Due to different demand sizes, vehicle capacities, etc. the schedule shown in 3 might be a likely result of solving four CARP problems independently. It could, however,

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12
General and organic	x	x	x	x	x	x	x	x	x	x	x	x
Paper and cardboard	x				x				x			
Glass and metal			x			x			x			x
Plastic		x		x		x		x		x		x

Figure 3.: Example of an inconvenient schedule for the citizen.

be very inconvenient for the citizen, in particular when the citizen is responsible for moving the waste bins from the back yard to the curbside and back. Consider now the schedule shown in Figure 4 in which the collection is coordinated to take place on the same weekday. The frequencies and the collection weeks have not been changed. Nonetheless, this schedule is significantly more attractive for the citizen than the one in Figure 3.

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12
General and organic	x	x	x	x	x	x	x	x	x	x	x	x
Paper and cardboard	x				x				x			x
Glass and metal			x			x			x			x
Plastic		x		x		x		x		x		x

Figure 4.: Example of a more convenient schedule for the citizen.

In this section, we define the Coordinated CARP (C-CARP) in which a number of individual CARPs with time perspectives are coordinated to provide service to the same edges on the same weekday. We consider the problem where each vehicle collects one waste fraction only.

The additional notation needed in this section is summarized in Table 2. In order to coordinate, we need to define a concept of a week. Define τ to be the number of days in a week. Let again $\mathcal{T} = \{1, 2, 3, \dots\}$ be the set of days in the planning horizon and define $\mathcal{T}_0 \subset \mathcal{T}$ to be the set of days on which service is not provided (e.g. weekends). \mathcal{T} and τ are further discussed below. Define subsets of \mathcal{T} as follows: $D_d = \{t \in \mathcal{T} | (t \bmod \tau) = d\} \cap \mathcal{T}_0$ for all $d = 1, \dots, \tau - 1$ and $D_\tau = \mathcal{T} \setminus \{\mathcal{T}_0 \cup \bigcup_{d=1}^{\tau-1} D_d\}$. These sets are sets of weekdays: D_1 contains all Mondays, D_2 Tuesdays, etc. and are sets used to enforce the coordination described above.

Let again \mathcal{F} be the set of all waste fractions to be collected and define for each $f \in \mathcal{F}$, the number of days between two consecutive collections of f on an edge as l^f . In order for full coordination to be possible, τ must be defined such that for every $f \in \mathcal{F}$, either $(l^f \bmod \tau) = 0$ or $(\tau \bmod l^f) = 0$. In other words, l^f must take values in $\{\dots, \frac{\tau}{3}, \frac{\tau}{2}, \tau, 2\tau, 3\tau, \dots\}$. We illustrate this with a few examples. Setting $\tau = 6$ gives us 6 working days a week (omitting e.g. Sundays). This is the natural choice in many cases, because it allows for service 3 times a week, twice a week, weekly, biweekly, every three weeks, etc. without violating above requirements. We can keep the same options for frequencies in a five-working day setting, by keeping $\tau = 6$, but including days of weekday 6 in \mathcal{T}_0 . If a 7-day workweek is needed, we can set $\tau = 8$ (operating with a 8 day week) and let \mathcal{T}_0 contain all weekday 8 days. This will allow service 4 times a week, twice a week, weekly, biweekly, every three weeks, etc. If some frequencies do not match the choice of τ , they can still be included by manually modifying the frequency constraints (28) below. This way of using \mathcal{T}_0 means that a fraction that is collected twice a week will have the four-day interval over the weekend if $\tau = 6$, and similarly for other combinations.

Define the subset $\tilde{\mathcal{F}} \subsetneq \mathcal{F}$ to be the set of waste fractions needing service multiple times per week, i.e. $l^f < \tau \forall f \in \tilde{\mathcal{F}}$. We refer to these as frequent waste fractions. The non-frequent fractions are given as $f \in \mathcal{F} \setminus \tilde{\mathcal{F}}$ and have $l^f \geq \tau$. In the coordination, we require the service of all non-frequent fractions to take place on the same weekday and we require that one of the weekly services of the frequent fractions also happens

Index Sets and Indices

$\tilde{\mathcal{F}}$	The subset of \mathcal{F} containing waste fractions needing service multiple times a week.
\mathcal{K}^f	The set of vehicles used to collect waste fraction $f \in \mathcal{F}$.
\mathcal{T}_0	The set of days on which service can not be performed.
D_d	The set of days defined as weekday d . $D_d \subseteq \mathcal{T}, d = 1, \dots, \tau$.

Parameters

W^k	Capacity of vehicle $k \in \mathcal{K}^f$ for fraction $f \in \mathcal{F}$.
l^f	The average number of days between two consecutive collections of $f \in \mathcal{F}$.
γ^k	Compression factor for waste fraction $f \in \mathcal{F}$ when collected by $k \in \mathcal{K}^f$.
τ	The number of days in a week.

Binary Variables

y_{ij}^{kt}	1 if vehicle $k \in \mathcal{K}^f$ performs service on $(i, j) \in \mathcal{E}_R^f$ at time $t \in \mathcal{T} \setminus \mathcal{T}_0$ and zero otherwise. Note that k only services one of the waste fractions.
---------------	---

Table 2.: Overview of additional notation for Coordinated CARP.

on that day.

After defining the week length τ and all the frequencies, the planning horizon \mathcal{T} must be determined such that the plan can be repeated. Let $|\mathcal{T}|$ equal the least common multiple of all l^f values. In that way, we have $(|\mathcal{T}| \bmod l^f) = 0 \forall f \in \mathcal{F}$, resulting in a cyclic plan.

The vehicles considered in this problem collect only a single waste fraction each. We therefore partition the set of vehicles, \mathcal{K} into $|\mathcal{F}|$ non-overlapping subsets such that \mathcal{K}^f is the set of vehicles used to collect waste fraction $f \in \mathcal{F}$. Thus once the vehicle is known, the waste fraction is given. The vehicles in each sets \mathcal{K}^f may or may not be identical. For each vehicle $k \in \mathcal{K}^f$, we define W^k to be the capacity and γ^k to be the compression factor. For real-world practical reasons (e.g. for the crew to feel ownership of their routes) we require the same vehicle to service a given fraction on a given edge throughout the planning horizon. We keep the time concept from section 3.1.3 and can therefore reuse the traversal and the vehicleused variables, but new service variables are needed. We define the following for every time $t \in \mathcal{T}$, every vehicle $k \in \mathcal{K}^f$ servicing waste fraction $f \in \mathcal{F}$, and every edge $(i, j) \in \mathcal{E}_R^f$ requiring service of that waste fraction.

$$y_{ij}^{kt} = \begin{cases} 1 & \text{if vehicle } k \in \mathcal{K}^f \text{ services edge } (i, j) \in \mathcal{E}_R \text{ at time } t \in \mathcal{T} \setminus \mathcal{T}_0, \\ 0 & \text{otherwise.} \end{cases}$$

The C-CARP can now be described as follows.

$$\text{lex min} \left(\sum_{k \in \mathcal{K}} C^k v^k, \sum_{k \in \mathcal{K}} \sum_{(i,j) \in \mathcal{E}} c_{ij} (x_{ij}^{kt} + x_{ji}^{kt}) \right)$$

$$\text{st.} \quad \sum_{k \in \mathcal{K}^f} \sum_{t=1, \dots, l^f} y_{ij}^{kt} = 1 \quad \forall f \in \mathcal{F}, (i,j) \in \mathcal{E}_R^f \quad (27)$$

$$y_{ij}^{kt} = y_{ij}^{k, t+l^f} \quad \forall f \in \mathcal{F}, k \in \mathcal{K}^f, \\ t = 1, \dots, |\mathcal{T}| - l^f, (i,j) \in \mathcal{E}_R^f \quad (28)$$

$$l^f \sum_{k \in \mathcal{K}^f} \sum_{t \in D_d} y_{ij}^{kt} = l^{f'} \sum_{k \in \mathcal{K}^{f'}} \sum_{t \in D_d} y_{ij}^{kt} \quad \forall f, f' \in \mathcal{F} \setminus \tilde{\mathcal{F}}, f \neq f', \\ d = 1, \dots, \tau, (i,j) \in \mathcal{E}_R^f \cap \mathcal{E}_R^{f'} \quad (29)$$

$$\sum_{k \in \mathcal{K}^f} \sum_{t \in D_d} y_{ij}^{kt} \leq \sum_{k \in \mathcal{K}^{\tilde{f}}} \sum_{t \in D_d} y_{ij}^{kt} \quad \forall f \in \mathcal{F}, \tilde{f} \in \tilde{\mathcal{F}}, \\ d = 1, \dots, \tau, (i,j) \in \mathcal{E}_R^f \cap \mathcal{E}_R^{\tilde{f}} \quad (30)$$

$$\sum_{(i,j) \in \mathcal{E}_R^f} q_{ij}^f y_{ij}^{kt} \leq \gamma^k W^k v^k \quad \forall f \in \mathcal{F}, k \in \mathcal{K}^f, t \in \mathcal{T} \quad (31)$$

$$(x_{ij}^{kt} + x_{ji}^{kt}) \geq y_{ij}^{kt} \quad \forall f \in \mathcal{F}, k \in \mathcal{K}^f, t \in \mathcal{T}, (i,j) \in \mathcal{E}_R^f \quad (32)$$

$$\text{Constraints (21) - (23)} \quad (33)$$

$$v^k \in \{0, 1\} \quad \forall f \in \mathcal{F}, \forall k \in \mathcal{K}^f \quad (34)$$

$$y_{ij}^{kt} = 0 \quad \forall f \in \mathcal{F}, k \in \mathcal{K}^f, t \in \mathcal{T}_0, (i,j) \in \mathcal{E}_R^f \quad (35)$$

$$y_{ij}^{kt} \in \{0, 1\} \quad \forall f \in \mathcal{F}, k \in \mathcal{K}^f, t \in \mathcal{T}, (i,j) \in \mathcal{E}_R^f \quad (36)$$

The objective function is the same as for the Multi-Day Commodity-Split MC-CARP. Constraints (27) ensure that every waste fraction $f \in \mathcal{F}$ is serviced within the first l^f days for every edge and (28) ensure that every subsequent service of the edge occurs with the correct frequency. The latter also ensures that the same vehicle collects waste fraction f from edge (i, j) every time collection of f is performed. (29) are the coordination constraints for non-frequent waste fractions. Consider a waste fraction $f \in \mathcal{F}$ and an edge (i, j) . Because τ divides l^f , (28) ensure that f is always collected from (i, j) on the same weekday d . For this weekday, the left hand side of (29) is the total number of collections of f from (i, j) during the time horizon times the frequency of f , i.e. the left hand side is $|\mathcal{T}|$ for weekday d . In order to make the right hand side, which is the same expression just for $f' \in \mathcal{F}$, $|\mathcal{T}|$ as well, the collections of f' must also take place on weekday d . As regards all other days, both sides of the equation are zero because no collections take place. Constraints (30) consider the frequent waste fractions that need service multiple times a week and ensure that at least one of those days is the same as the day on which the non-frequent waste fractions on the edge are serviced. On the weekday, d of service of the non-frequent waste fraction $f \in \mathcal{F} \setminus \tilde{\mathcal{F}}$, the left hand side of (30) gives the number of times f is serviced during the time horizon. By definition, this is no more than $|\mathcal{T}|/\tau$. The constraints state that any frequent waste fraction $\tilde{f} \in \tilde{\mathcal{F}}$ must also be serviced at least that many times on this

weekday. In fact, due to constraints (28), the right hand side will become $|\mathcal{T}|/\tau$ in this case, i.e. the frequent waste fractions will be serviced on day d in every week. For any other weekdays, the left hand side is zero, and the frequent fractions may be serviced or not depending on their frequency. Constraints (31) ensure that the vehicle capacities are respected and that nothing is serviced by vehicles that are not used in the solution. (32) force vehicles to traverse the edges they service. Finally, (33) are the route topology constraints inherited from the Multi-Day Commodity-Split MC-CARP model, and (34)-(36) define the domain for the service and vehicle-usage variables.

A natural extension of this problem would be to combine it with the No-Split MC-CARP such that service by a number of multi-compartment (and possibly single compartment) vehicles are to be coordinated over time. Another extension is to assign the frequency of a fraction for each edge individually. This would imply that the single fraction problems would become Periodic CARP instances instead of CARP instances as they would be above.

3.3. *Semi-Periodic CARP*

When considering problems such as waste collection, where the demand of an edge is the joint demand of the customers of a street segment, the service schedules of those customers may not be the same. Consider, for example, a street segment where three customers request collection once every two weeks (A), two customers every week (B), and one customer twice a week (C), and assume for the sake of simplicity that the demand of each customer is 1 unit. Then the demand of the edge will be 1 unit two times a week plus two additional units once a week, plus three additional units every second week. This is the problem considered in this section. It is referred to as the Semi-Periodic CARP (SP-CARP). Alternatively, the problem could be modeled as node routing, but by modeling it as arc routing, we preserve the concept of streets and avoid an increase in the number of required units. A similar problem is encountered in street sweeping applications, where the streets sometimes need more extensive cleaning than other times (Eglese and Murdock 1991).

For this problem, let again \mathcal{T} be the set of days in the planning horizon and \mathcal{T}_0 be the set of days on which service is not possible (e.g. weekends). As for C-CARP, we use τ to denote the number of days in a week. Define \mathcal{H} to be the set of different service schedules (i.e. 3 times a week, 2 times a week, weekly, bi-weekly, every 3 weeks, etc.). For each $h \in \mathcal{H}$, we let l^h be the average number of days between two consecutive collections of waste with service schedule h . In our example, $l^h = 14$ for A, $l^h = 7$ for B, and $l^h = 3.5$ for C. We define $a_h, b_h \in \mathbb{Z}_+$ to be the minimum and maximum number of days between two consecutive collections of waste with service schedule h . These parameters define the spacing of the service and are given as $a_h = \lfloor l^h \rfloor$ and $b_h = \lceil l^h \rceil$. In our example, this gives us $a_h = b_h = 14$ for A, $a_h = b_h = 7$ for B, and $a_h = 3, b_h = 4$ for C.

For each $h \in \mathcal{H}$, we define a cycle period as $\tilde{l}^h = \max\{\tau, l^h\}$. In our example we have $\tilde{l}^h = 14$ for A, $\tilde{l}^h = 7$ for B, and $\tilde{l}^h = 7$ for C. The cycle period is repeated throughout the planning period as regards the demand to be serviced on each street segment and the vehicle to perform the service. This means that for service schedules with multiple services per week, the weekly schedule will be repeated whereas less frequent service schedules are only repeated as needed.

We use a homogeneous set of vehicles \mathcal{K} , each with capacity W . As there is a time aspect like in Multi-Day Commodity-Split MC-CARP, we use v^k to denote whether a

vehicle is included or not. Let q_{ij}^h be the demand of edge (i, j) with service schedule $h \in \mathcal{H}$. For each $h \in \mathcal{H}$, we define $\mathcal{E}_R^h = \{(i, j) \in \mathcal{E} : q_{ij}^h > 0\}$ and set $\mathcal{E}_R = \bigcup_{h \in \mathcal{H}} \mathcal{E}_R^h$. The traversal variables are unchanged compared to the previous models with a time concept, and the service variables are defined $\forall t \in \mathcal{T}, \forall k \in \mathcal{K}, \forall h \in \mathcal{H}, \forall (i, j) \in \mathcal{E}_R^h$ as follows.

$$y_{ij}^{hkt} = \begin{cases} 1 & \text{if the demand of service schedule } h \in \mathcal{H} \text{ of edge } (i, j) \in \mathcal{E}_R^h \\ & \text{is serviced by vehicle } k \in \mathcal{K} \text{ at time } t \in \mathcal{T} \setminus \mathcal{T}_0, \\ 0 & \text{otherwise.} \end{cases}$$

Note that y_{ij}^{hkt} is also defined for $t \in \mathcal{T}_0$, but that the below constraints (44) ensure that no service can be performed on these days. The SP-CARP can now be modeled as follows.

$$\text{lex min} \left(\sum_{k \in \mathcal{K}} C^k v^k, \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \sum_{(i, j) \in \mathcal{E}} c_{ij} (x_{ij}^{kt} + x_{ji}^{kt}) \right)$$

$$\text{st.} \sum_{k \in \mathcal{K}} \sum_{t=1, \dots, \tilde{t}^h} y_{ij}^{hkt} = \max\{1, \frac{\tau}{\tilde{t}^h}\} \quad \forall h \in \mathcal{H}, (i, j) \in \mathcal{E}_R^h \quad (37)$$

$$\sum_{k \in \mathcal{K}} y_{ij}^{hkt} = \sum_{k \in \mathcal{K}} y_{ij}^{hkt + \tilde{t}^h} \quad \forall h \in \mathcal{H}, t = 1, \dots, |\mathcal{T}| - \tilde{t}^h, (i, j) \in \mathcal{E}_R^h \quad (38)$$

$$\sum_{k \in \mathcal{K}} \sum_{t'=t, \dots, t+b_h} y_{ij}^{hkt'} \geq 1 \quad \forall h \in \mathcal{H}, t = 1, \dots, |\mathcal{T}| - b_h, (i, j) \in \mathcal{E}_R^h \quad (39)$$

$$\sum_{k \in \mathcal{K}} \sum_{t'=t, \dots, t+a_h} y_{ij}^{hkt'} \leq 1 \quad \forall h \in \mathcal{H}, t = 1, \dots, |\mathcal{T}| - a_h, (i, j) \in \mathcal{E}_R^h \quad (40)$$

$$\sum_{h \in \mathcal{H}} \sum_{(i, j) \in \mathcal{E}_R^h} q_{ij}^h y_{ij}^{hkt} \leq \gamma W v^k \quad \forall k \in \mathcal{K}, t \in \mathcal{T} \quad (41)$$

$$|\mathcal{H}|(x_{ij}^{kt} + x_{ji}^{kt}) \geq \sum_{h \in \mathcal{H}: q_{ij}^h > 0} y_{ij}^{hkt} \quad \forall k \in \mathcal{K}, t \in \mathcal{T}, (i, j) \in \mathcal{E}_R, \quad (42)$$

$$\text{Constraints (21) – (23)} \quad (43)$$

$$y_{ij}^{hkt} = 0 \quad \forall h \in \mathcal{H}, k \in \mathcal{K}, t \in \mathcal{T}_0, (i, j) \in \mathcal{E}_R^h \quad (44)$$

$$y_{ij}^{hkt} \in \{0, 1\} \quad \forall h \in \mathcal{H}, k \in \mathcal{K}, t \in \mathcal{T} \setminus \mathcal{T}_0, (i, j) \in \mathcal{E}_R^h \quad (45)$$

$$v^k \in \{0, 1\} \quad \forall k \in \mathcal{K}. \quad (46)$$

Constraints (37) ensure the correct number of services in the first cycle period for each service schedule h and (38) ensure that the pattern is repeated throughout the time horizon. (39) and (40) are the spacing constraints. Here, (39) ensure that requests with service schedule $h \in \mathcal{H}$ are serviced at least every b_h days, and (40) ensure that they are serviced at most every a_h day. (41) ensure that only vehicles used in the solution are assigned demand and that vehicle capacity is respected. (42) ensure that edges can only be serviced if they are traversed. Finally, (43) are the route topology constraints and (44)-(46) define the domain of the service and vehicle used variables.

The model ensures cyclic planning when the plan is repeated as long as τ and l^h divide $|\mathcal{T}|$. If the problem at hand does not contain the concept of a week, the above model can still be used with some modifications. In that case, we set $\tau = \mathcal{T}$. The spacing parameters a_h and b_h can take any value $b_h \geq a_h$ to incorporate more flexibility in the model, and the following two constraints must be added to ensure that planning is periodic when the plan of the model is repeated.

$$\sum_{k \in \mathcal{K}} \left(\sum_{t'=t, \dots, |\mathcal{T}|} y_{ij}^{hkt'} + \sum_{t'=1, \dots, b_h - (|\mathcal{T}| - t + 1)} y_{ij}^{hkt'} \right) \geq 1 \quad \forall h \in \mathcal{H}, (i, j) \in \mathcal{E}_R^h$$

$$t = |\mathcal{T}| - b^h + 1, \dots, |\mathcal{T}| \quad (47)$$

$$\sum_{k \in \mathcal{K}} \left(\sum_{t'=t, \dots, |\mathcal{T}|} y_{ij}^{hkt'} + \sum_{t'=1, \dots, a_h - (|\mathcal{T}| - t + 1)} y_{ij}^{hkt'} \right) \leq 1 \quad \forall h \in \mathcal{H}, (i, j) \in \mathcal{E}_R^h$$

$$t = |\mathcal{T}| - a^h + 1, \dots, |\mathcal{T}| \quad (48)$$

Index Sets and Indices	
\mathcal{E}_R^h	The set of required edges wrt. $h \in \mathcal{H}$, i.e. $\mathcal{E}_R^h = \{(i, j) \in \mathcal{E} : q_{ij}^h > 0\}$.
\mathcal{H}	The set of service schedules, indexed by h .
Parameters	
a_h	Minimum number of days between two consecutive services of $h \in \mathcal{H}$ on the same edge.
b_h	Maximum number of days between two consecutive services of $h \in \mathcal{H}$ on the same edge.
l^h	The average number of days between two consecutive collections of waste with service schedule $h \in \mathcal{H}$.
\bar{l}^h	$\max\{\tau, l^h\}$. The cycle period for service schedule $h \in \mathcal{H}$.
q_{ij}^h	Demand of edge $(i, j) \in \mathcal{E}_R$ with service schedule $h \in \mathcal{H}$.
Binary Variables	
y_{ij}^{hkt}	1 if the demand of service schedule $h \in \mathcal{H}$ of edge $(i, j) \in \mathcal{E}_R$ is serviced by vehicle $k \in \mathcal{K}$ at time $t \in \mathcal{T} \setminus \mathcal{T}_0$, and zero otherwise.

Table 3.: Overview of additional notation for the Semi-Periodic CARP.

In the SP-CARP described above, the demand to be collected remains the same regardless of the duration between two consecutive services of an edge. This assumption can be justified by the fact that our collaborators currently have a fixed duration between service times and thereby fixed demand size. However, a natural extension of this problem is to let the demand be time dependent as is the case in Dos Santos et al. (2016). Furthermore, the semi-periodic nature of the demand studied here could naturally be included in the MC-CARP and C-CARP variations.

4. New Large-Scale Benchmark Data

Five new CARP variations were introduced in the previous section and in this section we present large-scale benchmark instances for each of these problems as well as for the CARP. All new data sets are available at <http://www.optimization.dk>, which also contains electronic appendices showing characteristics of the data. In Section 4.1, we describe the process of obtaining graphs for instances based on our raw data and in Section 4.2, we briefly explain our use of the terminology of graphs, vehicle files, and instances. The remaining four sections are devoted to the different CARP variations presented in the paper and to a description of important characteristics of our new benchmark instances.

The large-scale data presented in this paper is based on real-life road networks as well as real-life demand from our five collaborators. To illustrate the structural differences of the data of the five areas, Figure 5 shows the complete network of each. Based on each of these five areas, we have created a number of graphs. This is further described in Section 4.1. The figure clearly shows that the F-data has traditional city center structure, in the O-data the city-and-suburb structure is dominating, whereas the data related to K, N, and S are mainly rural in nature. The odd shape of the S-data is mainly due to the sea, and the K-data represents a combination of two counties.

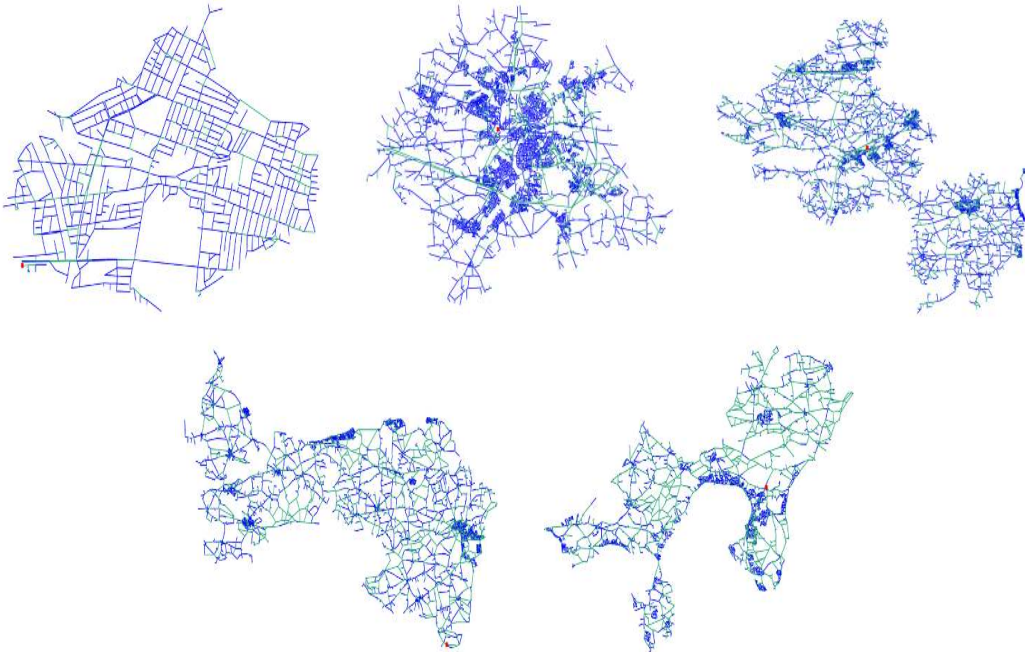


Figure 5.: The five data areas. From the top left: Frederiksberg (F), Odense (O), Skanderborg and Odder (K), North Djurs (N), and South Djurs (S). Blue edges are required and green edges are non-required. The depot is represented by a red square.

4.1. Data Treatment

This section is devoted to a short description of the way the original data was treated in order to obtain the final graphs. Section 4.1.1 describes how the base graphs are created and explains the naming convention. In Section 4.1.2, we describe how the

demand was obtained and Section 4.1.3 outlines the graph structural modifications.

4.1.1. *Creating the Graphs*

In order to create the base graphs, i.e. the graphs underlying all our instances, we received two types of files from KortCenter A/S (2017). The street file contained the coordinates of the end points as well as the length of each street segment in meters together with street name and zip code information. Based on this, we created a network for each of the five areas. The household file contained the coordinates for each household as well as complete address information.

We combined these files by using the address information to identify the subset of street segments that might be relevant for each household, and then we projected the household onto the nearest of these segments. After this process, each edge of the network contained address information on all households assigned to it.

The next step in the process was to create the base graphs. All base graphs numbered 1 through 9, are based on the complete network as described above, and they differ only in the demand assignment as described in Section 4.1.2. In order to also obtain instances of smaller size, we created base graphs by selecting a center node and a radius. Based on these, all nodes with a shortest path distance to the center node less than the radius and the edges between them were gathered to create a base graph. We manually selected two center nodes (instance numbers 10-13 and 14-17, respectively) for each area. We then selected radius' for each center node such that the number of nodes would be approximately 50% (instance numbers 10 and 14), 25% (instance numbers 11 and 15), 10% (instance numbers 12 and 16), and 3% (instance numbers 13 and 17) of the original number of nodes. The depot node is the original depot location when included in the graph; when that is not the case, the depot is chosen to be the center node used to create the smaller graph.

The naming of the instances is illustrated by the example CARP-F1_g, where the 'F' refers to the area of Frederiksberg. Alternative values are 'K', 'N', 'O', and 'S'. The 'g' at the end of the name refers to the demand which originates from the mixed general waste fraction. Alternatively, 'p' is used when the demand originates from paper. This is described in the next section. Finally, index numbers are as just described.

4.1.2. *Demand*

In order to assign demand to the edges of the graphs, we received a waste file for each of the five areas from our collaborators. These files contained address information and information about the waste fractions to be collected from each individual household, the bin sizes, and the collection schedules. This information was coupled to the edges of the base graphs using the address information. The edges of the base graphs thereby hold non-aggregated and detailed demand information.

In addition, we received estimates regarding the historical average fill levels of the bins, which we used to down-scale all amounts slightly compared to the bin sizes. Furthermore, a few edges had very high demand originating from container-like bins. We have down-scaled them further manually, but they remain among the largest demands. This was done as they caused infeasibility issues, in particular for the separated waste because the vehicle compartments are smaller.

Four of our five areas only separately collect paper and general (mixed) waste from the households. In order to create instances for multiple separated fractions, we have used information from Econet AS (2012) to determine the mix of the general waste.

The report states the mix according to weight (not volume). Figure 6 shows the mix when converted into volume and it forms the basis of our separation of waste for the MC-CARP and C-CARP instances. In the F area there is a higher degree of sorting at the households, and we have therefore not performed this separation, but instead used their demand information directly.

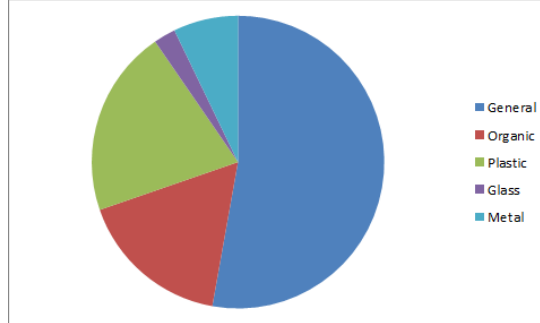


Figure 6.: Mix of the general waste based on volume. Source: Econet AS (2012) in a slightly modified form.

For the SP-CARP, we have used the actual service schedules for each edge, whereas for all variations of MC-CARP as well as for the CARP, we have normalized the demand to match a two-week schedule. For the C-CARP, the demand has been normalized to the collection intervals of each set.

Three categories of households are present in our areas: Apartment buildings (A), residential houses (R), and summer residences (S and s). Summer residences often require service less frequently in the winter (s) than in the summer (S). All instances based on subgraphs of the complete networks (10-17) use demand from all households and summer schedules (ARS). For the large graphs (1-9), different demand combinations are used. For instance, graph N4_g uses demand from regular houses and summer residences with winter schedule (Rs). Details can be found for CARP and SP-CARP in the electronic appendices. For C-CARP and the MC-CARP variations, only graph numbers 1, 10-17, all using ARS demand are used. Finally, areas F and O do not contain any summer residences, which causes a reduction in the number of graphs.

4.1.3. Structural Changes

The resulting graphs were then modified such that unwanted attributes were removed. We did so to avoid edges that would never be used in an optimal solution and to provide well-structured graphs for future algorithms. Below, the method used to remove each of the attributes is described in the order in which the changes were done. The first two modifications remove unnecessary deadheading edges whereas the purpose of the last two are purely to ease the use of the graphs for solution algorithms. None of the modifications have any effect on the optimal solutions.

To simplify the description, we denote edges as $e = (i, j)$ where i and j are unordered, and we use q_e to denote any kind of demand on e . Let $\delta(i)$ be the degree of node $i \in \mathcal{N}$ and $\delta_R(i)$ the degree with respect to required edges.

4.1.3.1. Remove empty areas.: Knowing that some parts of the graphs are without demand, we first remove those parts that can be removed without altering the

optimal solution. This is done as follows. For each node $i \in \mathcal{N}$, initialize

$$\rho(i) = \begin{cases} 1 & \text{if } i \text{ is adjacent to at least one required edge,} \\ 0 & \text{otherwise.} \end{cases}$$

Next, based on pre-calculated shortest path matrices, set $\rho(i) = 1$ for all nodes on the shortest path between any pair of nodes adjacent to a demand edge. Finally, we delete all nodes $i \in \mathcal{N}$ where $\rho(i) = 0$ from the graph along with any edges adjacent to node i . This is illustrated in Figure 7.

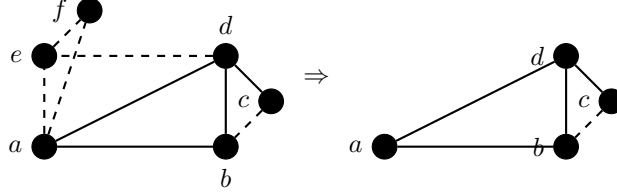


Figure 7.: Remove empty areas. Solid and dashed lines represent required and non-required edges, respectively.

4.1.3.2. Shorten paths.: A node of degree two, where both adjacent edges are in $\mathcal{E} \setminus \mathcal{E}_R$, will only be traversed as part of a shortest path between two consecutive services. Therefore, the adjacent edges will always be traversed together or not at all. Hence, the node can be removed and the two adjacent edges can be concatenated. Let i be such a node and let $e = (i, j)$ and $e' = (j, i')$ be the two adjacent edges in \mathcal{E}_R . We delete j , e , and e' from the graph, and add a new edge $\hat{e} = (i, i')$ with $c_{\hat{e}} = c_e + c_{e'}$ and $q_{\hat{e}} = 0$. This continues until no such adjacent edges exist. This is illustrated in Figure 8.

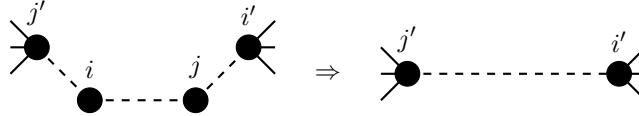


Figure 8.: Shorten paths. Solid and dashed lines represent required and non-required edges, respectively.

4.1.3.3. Remove loops.: In order to make the graphs easier to work with, we first modify them by removing loops $e = (i, i) \in \mathcal{E}$. We have to two different cases to consider. Either the loop has no demand, $q_e = 0$, in which case the loop will simply be removed from the graph. Or we have a loop with demand $q_e > 0$. In this case, the loop cannot be removed. Instead, a new node is added somewhere on the edge. This modification is illustrated in the left part of Figure 9. We create a new node j and add this to the graph. The edge $e = (i, i)$ is modified such that one of its endpoints is the new node, i.e. $e = (i, j)$ and the cost is reduced slightly $c_e = c_e - 1$. The demand of e is unchanged. A new edge is added to the graph $e' = (j, i)$ with $c_{e'} = 1$ and $q_{e'} = 0$. As the costs of every edge in the original graph are larger than 2, this creates a pair of parallel edges of which the longest is required. Parallel edges are another unwanted attribute handled in the procedure below.

4.1.3.4. Remove parallel edges: The last change is removal of parallel edges $e, e' \in \mathcal{E}$ where $e = (i, j)$ and $e' = (i, j)$ for $i, j \in \mathcal{N}$. Again two variations should be considered. First, we have the scenario where the longest of the parallel edges is not required. In this case, the edge can be removed as an optimal solution will always choose to traverse the other edge when traversing between i and j . In the other case, we must keep both edges as the longest of them has to be traversed by the route servicing it, and the shorter is most favorable for deadheading. This case is illustrated at the right hand side of Figure 9. Assume that $c_e < c_{e'}$ and $q_{e'} > 0$. A new node i' is added to the graph and the longest edge e' is modified $e' = (i', j)$ with cost $c_{e'} = c_e - 1$. Furthermore, a new edge $\hat{e} = (i, i')$ is added to the graph with cost $c_{\hat{e}} = 1$ and demand $q_{\hat{e}} = 0$. The shortest edge e is not changed.

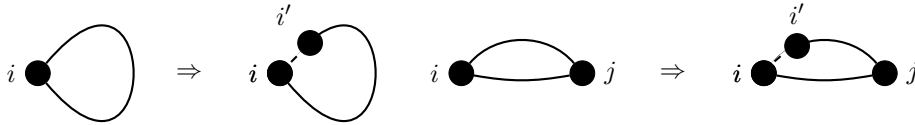


Figure 9.: Left: remove loops. Right: Remove parallel edges. Solid and dashed lines represent required and non-required edges, respectively.

4.2. Graphs, Vehicles, and Instances

We have based our benchmark instances for the five new problems and our new benchmark instances for the CARP on a special terminology that defines the terms 'graph' and 'instance' in the following way.

Graphs contain all the information related to the demand side of the problems: The network, the depot, and cost information as well as the demand of each edge at the level of detail needed for the problem, be it one or multiple fractions, normalized or non-normalized. They also provide information on the number of waste bins at each edge even though that information is not strictly necessary for the problems presented in this paper. Furthermore, each graph file contains information regarding the origin of the data; the county, radius, types of houses, etc.

Vehicle files contain information related to the service side of the problems: The number of different vehicle types, and for each type, the number of available vehicles and their number of compartments. They also contain information regarding compartment sizes and compression factors. Furthermore, the vehicle files contain information regarding \mathcal{T} , \mathcal{T}_0 , and τ for instances where these are relevant.

Finally, the vehicle files are prepared such as to provide information to be used for inclusion of time duration constraints or a maximum number of bins per route. In the present form, these parameters are given dummy values, but the standardized syntax including that information eases future research.

We use the term 'instance' to denote a combination of a graph and a vehicle file. Thereby, each graph serve as the foundation for several instances. The separation of the two eases any kind of analysis where the effect of servicing a graph (network and demand) with different types of vehicles is to be studied.

	New large-scale instances					Classical benchmark instances				
	F	K	N	O	S	EGL	GDB	KSHS	VAL	BE
Number of instances	42	60	60	42	60	24	23	6	34	10
Number of graphs	14	20	20	14	20	24	23	6	34	10
Min number of nodes	26	375	268	228	322	77	7	6	24	255
Max number of nodes	812	11640	8537	10283	6149	140	27	10	50	255
Min number of edges	33	412	305	247	374	98	11	15	34	375
Max number of edges	1124	12675	9725	11863	7110	190	55	15	97	375
Av. percentage of edges being required	62.9	59.3	58.9	65.2	47.0	76.9	100	100	100	96.3
Av. # of req. edges per demand component	8.0	3.0	3.4	6.0	2.2	58.3	29.4	15	63.3	361
Av. node degree	2.7	2.3	2.3	2.4	2.4	2.6	5.0	3.8	3.5	2.9
Av node degree wrt. required edges	1.7	1.3	1.4	1.6	1.1	2.0	5.0	3.8	3.5	2.8
Av. Percentage of nodes with degree 1	11.2	27.5	22.1	26.8	23.0	16.0	1.1	3.3	0	6.3
Av. Percentage of nodes with degree 2	16.6	23.4	27.1	13.8	19.5	19.8	5.3	11.6	16.4	10.2

Table 4.: Summary of the CARP graphs.

Throughout the assignment of vehicles, we have sought to use realistic vehicles and meaningful compression factors for each fraction. However, in order to obtain interesting instances for multiple fractions that do not simply reduce to single commodity instances, we have adjusted compartment size and compression to the limit of what would be realistic. Furthermore, the smallest instances in particular have been assigned very small vehicles to ensure that they require more than a single vehicle to be used.

4.3. Data for the CARP

Based on the five areas, we have created five sets of data, containing a total of 88 graphs. Each graph is associated with three different vehicles, which gives us 264 CARP instances. Table 4 provides a short summary of the graphs and Appendix ?? provides details of each graph together with the vehicles associated with it. In the summary table, we have provided the corresponding characteristics for the four sets of classical benchmark instances for comparison: EGL (Li and Eglese 1996), GDB (Golden et al. 1983), KSHS (Kiushi et al. 1995), VAL (Benavent et al. 1992), and BE (Brandão and Eglese 2008).

The majority of the new graphs are huge, containing thousands of nodes, but some are smaller. This is particularly the case for the F instances that originate from a small county. 243 (210) of the 264 new instances are larger than the largest existing benchmark instances when considering the number of nodes (required edges), and 177 (138) of the instances contain more than 1000 nodes (required edges). This illustrates the large-scale of the new data.

When we look across the five sets of instances, it is particularly apparent that in the two sets that originate from cities, O and particularly F, the percentage of the edges being required is significantly higher than in the more rural areas where only about half of the edges are required. For comparison, in the GDB, KSHS, and VAL instances, all edges are required and the EGL instances lie somewhere in-between. We also note the low number of nodes in each connected component when only required edges are considered. The last part of the table treat the node degree. Compared to the GDB, KSHS, and VAL instances, the average node degree, in particular with respect to required edges, is significantly smaller in the new instances. In fact, 20-25 percent of the nodes (less for the F instances) are deadends. Furthermore, the graphs contain a significant amount of nodes with degree 2. Since the graphs have already been cleaned up as described in Section 4.1, at least one of the adjacent edges of these nodes is

required. For comparison, the existing benchmark instances are much more dense and many of them contain only one connected component when considering required edges. This is particularly the case for those instances that have been constructed rather than originating from real-life street networks.

4.4. Data for Multi-Commodity CARP

We used 5 graphs from each area and combined these with different demand compositions to create a total of 260 graphs. We then combined graphs and vehicles in reasonable ways and created 400 instances for the No-Split MC-CARP, 800 for the C-Split MC-CARP, and 464 instances for the Multi-Day C-Split MC-CARP. We explain the procedure in two steps below. First, we focus on the demand in each subset and on the resulting graphs and then we focus on the assignment of vehicle files, which altogether provide instances for the three variations of the problem. Details of both graphs and instances are given in the electronic appendix related to MC-CARP. We use the term MC-CARP to cover all three variations.

Set	General	Organic	Plastic	Metal	Glass	Paper	Cardboard
A	Mixed					X	
B	Mixed		Mixed			X	
C			X	X	X	X	
D	X	X	Mixed			X	
E	X	X	X	X	X	X	
G	X	X	X	X	X	X	X

Table 5.: Overview of the fractions included in each set of MC-CARP instances.

Four of our five areas only collect paper and general (mixed) waste separately from the households. In the F area, there is a higher degree of sorting at the households. As our study was initiated by the fact that a higher degree of recycling and thereby sorting is desired, we have created the MC-CARP instances based on a number of likely sorting scenarios. Table 5 provides an overview of the waste fractions included in each set. Each set contains 25 graphs (except set G). Set A is the current setup with a separate paper fraction and a mixed general fraction containing everything else. As the general fraction is sent for incineration in this case, high compression can be used. In set B, the so-called waste resources (plastic, glass, and metal) are separated from the general waste with the purpose of subsequent sorting. The resources can only be compressed slightly to enable sorting, whereas the general waste can still be compressed maximally. Skipping to set D, the general waste has been separated into organic waste and general waste. The data instances are created such that they allow comparison of these three scenarios. This is explained in more detail in Section 5. Set C contains the four fractions that are directly reusable (paper, plastic, glass, and metal), which in some areas (not contained in our data) are collected by a 4-compartment vehicle. Finally, in set E, all six fractions are separated, and set G includes the additional fraction (cardboard) currently collected in area F, thereby containing seven separate fractions.

Since only two fractions are currently being collected separately (except in area F), we have used Econet AS (2012) to obtain an estimate of the composition of the mixed general waste as explained in Section 4.1. Subsequently, because citizens are not

	A	R-A	B	R-B	C	R-C	D	R-D	E	R-E	G	R-G
# instances for No-Split MC-CARP	50	50	50	50	50	50	50	50				
# instances for C-Split MC-CARP			100	100			100	100	100	100	100	100
# instances for Multi-Day C-Split MC-CARP			72	72			72	72	72	72	16	16
# graphs	25	25	25	25	25	25	25	25	25	25	5	5
# fractions	2	2	3	3	4	4	4	4	6	6	7	7
Av. number of req. fractions per edge	1.1	0.9	1.8	1.6	2.3	1.7	2.3	1.9	3.5	2.6	3.1	2.3
Av. number of req. fractions per req. edge	1.8	1.4	2.7	2.4	3.6	2.7	3.6	3.0	5.4	4.1	4.8	3.6
Av. percentage of edges not req. service	36.2	36.2	36.2	36.3	36.5	37.0	36.2	36.2	36.2	36.2	35.1	35.1
Av. percentage of edges req. 1 fraction	14.0	26.7	1.2	3.8	1.2	6.7	1.2	4.3	1.2	1.9	4.8	6.2
Av. percentage of edges req. 2 fractions	49.8	37.1	13.7	27.3	0.4	18.5	0.6	13.9	0.2	4.3	0.8	7.2
Av. percentage of edges req. 3 fractions			48.9	32.6	22.7	26.0	17.9	28.0	0.8	12.2	2.9	14.9
Av. percentage of edges req. 4 fractions					39.2	12.4	44.1	18.4	4.3	20.5	3.5	20.8
Av. percentage of edges req. 5 fractions									18.8	18.0	40.3	12.4
Av. percentage of edges req. 6 fractions									38.5	6.8	10.2	3.3
Av. percentage of edges req. 7 fractions											2.3	0.1
Av. percentage of edges req. general			63.1	58.0			63.1	47.4	63.1	47.4	62.3	46.7
Av. percentage of edges req. organic							57.3	43.1	57.3	43.1	33.2	26.1
Av. percentage of edges req. plastic	63.6	63.4			62.2	46.8			61.9	46.4	56.0	42.8
Av. percentage of edges req. metal			62.2	60.7	62.1	46.9	62.2	60.8	61.8	46.5	55.5	42.0
Av. percentage of edges req. glass					52.2	39.3			52.2	39.2	7.7	5.4
Av. percentage of edges req. paper	50.0	37.5	50.0	37.4	50.4	37.9	50.0	37.5	50.0	37.5	59.4	44.2
Av. percentage of edges req. cardboard											34.2	23.9

Table 6.: Summary of the MC-CARP graphs.

perfect at sorting and in order to obtain true multi-commodity instances, a random amount of each of the waste fractions plastic, glass, metal, and organic at each edge is moved from that fraction to general waste. The procedure leads to the A, B, C, D, and E instances for areas K, N, O, and S. For the F area, the data is directly available, and we have not used this process.

To provide instances with even more variation, we have created an 'R-' version of each graph as follows. Based on each E instance, we have, for each edge, deleted the full amount of each fraction with a probability of 25 percent. However, in the process, we have ensured that not all waste fractions are deleted, i.e. required edges are still required. All these deletions are subsequently adapted to the R-A, R-B, R-C, and R-D instances. This ensures that comparison is still possible across the R-A, R-B, and R-D instances and that required edges remain the same. As the C instances is the only set that do not contain all fractions this modification imply that all required edges in the C instances are not necessarily required in the R-C instances.

Table 6 summarizes the characteristics of the 12 resulting sets, each containing 25 graphs (except G and R-G, each containing 5 graphs). The center part of the table provides information on the average number of fractions required per edge and about the average percentage of the edges requiring a given number of fractions, whereas the last part of the table states the average percentage of edges requiring each fraction. The last part of the table is structured in the same way as Table 5. The table clearly shows that the 'R-' instances contain less demand than the original ones. Details of the graphs are given in Tables 1 through 6 in the electronic appendix related to MC-CARP.

Because it is not realistic to have waste collection vehicles with more than four compartments, only sets A-D have been used for No-Split MC-CARP. For each of these graphs, we have assigned two vehicle files and thereby obtained a total of 400 No-Split MC-CARP instances. Details of the assignment can be seen in Tables 7 and 8 in the appendix.

For the C-Split MC-CARP and the Multi-Day C-Split MC-CARP, we have considered graphs where all waste is collected and where it is separated into at least three bins, i.e. sets B, D, E, and G. For both variations, each graph is combined with four

vehicle files, resulting in 800 C-Split MC-CARP instances and 464 Multi-Day C-Split MC-CARP instances after removal of irrelevant instances. The instances are detailed in Tables 9-10 and Tables 11-14 in the appendix, respectively.

4.5. *Data for Coordinated CARP*

	A	B	C	D	E
Number of instances	43	39	39	39	39
Number of graphs	25	25	25	25	25
Time horizon, days	42	42	36	84	12
Week length	7	7	6	7	6
Number of intervals	2	3	3	4	4
Intervals	14, 21	7, 14, 21	3, 12, 18	7, 14, 21, 28	2, 3, 6, 12
Number of service days	30	30	30	60	10
Av. percentage of edges not req. service	36.2	36.2	36.2	36.2	36.2
Av. percentage of edges req. 1 fraction	14.0	1.2	1.2	1.2	1.2
Av. percentage of edges req. 2 fractions	49.8	13.7	13.7	0.6	0.6
Av. percentage of edges req. 3 fractions		48.9	48.9	17.9	17.9
Av. percentage of edges req. 4 fractions				44.1	44.1

Table 7.: Summary of the C-CARP graphs

For the C-CARP, we have created five sets of instances, containing a total of 125 graphs and 199 instances. Set A contains two fractions to be collected every two and every three weeks, respectively. Sets B and C have three fractions to be coordinated. They use the same graphs and demand data, but differ in the service intervals. Finally, sets D and E coordinate four fractions. Again, they use the same graphs and demand data, but differ in the service intervals. Sets C and E contain frequent fractions, whereas sets A, B, and D do not. The instances with frequent fractions use a week length of six days with one weekly non-service day, whereas the other instances use seven day weeks with two non-service days. Hence, all instances have five weekly service days. A summary of the instances is given in Table 7 and complete information about the instances can be found in the electronic appendix related to C-CARP.

Each graph has been associated with either one or two vehicle files. In the first file, only one type of vehicles is available for each fraction, whereas in the second file, two types are available for each fraction, and the larger of the two is only available in limited amounts. In the smaller instances, the demand is so small that the second type is not a reasonable choice. Therefore, these graphs only have one vehicle file assigned.

4.6. *Data for Semi-Periodic CARP*

For the SP-CARP, we have created a total of 249 instances based on 83 graphs. We have used the same base graphs as for the CARP, and the slight reduction in the number of graphs for the SP-CARP is caused by the fact that some graphs turned out to have only a single collection interval, making them in fact a CARP instance. The instances are partitioned into five sets, one for each area. Table 8 provides a summary of the five sets, whereas details of the graphs and instances are found in the electronic appendix related to SP-CARP.

All demand intervals are directly adapted from the original service schedule of each household associated with each street segment of each area. Therefore, not all collection

intervals are present in every set. In Table 8, we use a '-' to indicate when an interval (or a combination of intervals) is not present at all, whereas a percentage of 0.0 indicates that the availability is too small to be noticed with one digit of precision. It is clear from the table that areas N, O, and S are relatively streamlined in their collection intervals, whereas K and in particular F are much more flexible. For each graph, the available service intervals jointly determine the time horizon. Each graph has been associated with three vehicles, where the difference lies in the capacity of the vehicle after compression.

The instances provided here assume a fixed collection interval. If the interval between collections should be decided by the model (periodic dependent demand) instead each demand should be divided by l^h so that demand per day would be $\frac{q_{ij}^h}{l^h}$.

	F	K	N	O	S
Number of instances	42	57	54	39	57
Number of graphs	14	19	18	13	19
Min number of nodes	26	375	268	228	322
Max number of nodes	812	11640	8537	10283	6149
Min number of edges	33	412	305	247	374
Max number of edges	1124	12675	9725	11863	7110
Av. percentage of edges being required	62.9	60.4	59.9	64.7	47.8
Av. percentage of edges req. service 3 times a week	8.6	0.0	-	-	-
Av. percentage of edges req. service 2 times a week	25.9	0.3	0.1	0.2	0.1
Av. percentage of edges req. service weekly	34.8	5.2	29.4	1.9	29.3
Av. percentage of edges req. service every 2 weeks	14.2	50.8	44.2	50.7	29.8
Av. percentage of edges req. service every 3 weeks	0.6	-	-	-	-
Av. percentage of edges req. service every 4 weeks	45.5	5.3	4.5	13.2	5.1
Av. percentage of edges req. service every 8 weeks	20.9	2.5	-	-	-
Av. percentage of edges not req. service	37.1	39.6	40.1	35.3	52.2
Av. percentage of edges req. 1 service interval	15.8	56.8	41.6	63.5	31.3
Av. percentage of edges req. 2 service intervals	16.9	3.5	18.3	1.2	16.5
Av. percentage of edges req. 3 service intervals	21.7	0.1	0.0	0.0	0.0
Av. percentage of edges req. 4 service intervals	7.1	-	-	-	-
Av. percentage of edges req. 5 service intervals	1.3	-	-	-	-
Av. percentage of edges req. 6 service intervals	0.2	-	-	-	-
Longest time horizon, weeks	8	8	4	4	4

Table 8.: Summary of the SP-CARP graphs.

5. Comparisons and Scenario Analysis

From a waste collection point of view, several scenario comparisons within and among the instances are interesting. We have had this in mind throughout the data generation process and the options are described below.

The first and most obvious question is whether to make separate or joint collection from all types of households. As apartment blocks generate waste much faster than residential or summer houses, they often have different types of bins that have to be handled by specialized vehicles. This difference makes an analysis of joint or separate collection an interesting issue from a waste point of view. This can be analyzed based on the CARP graphs by comparing the solution of an instance containing all three

types (ARS), e.g. S1_g to the two solutions obtained by separating the apartment blocks. In this case, it would correspond to the sum of the costs of S9_g and S3_g. This is illustrated in the first row of Table 9. In total, there are seven such possibilities for scenario comparison for each of the areas K, N, and S as shown in the table, whereas there is only one option for each of the F and O sets: AR versus A and R. In total, this gives 23 comparison options within the CARP as regards the graphs. For each option, vehicles should be assigned depending on the scenario to be analyzed.

Within the SP-CARP, we can perform the same kind of scenario analysis for the same combinations of household types, leading again to 23 comparison options within the SP-CARP.

Household types		Example files	
Scenario 1	Scenario 2	Scenario 1	Scenario 2
ARS	RS + A	S1_g	S3_g + S9_g
ARs	Rs + A	S2_g	S4_g + S9_g
ARS	AR + S	S1_g	S5_g + S7_g
ARS	A + R + S	S1_g	S9_g + S6_g + S7_g
AR	A + R	S5_g	S9_g + S6_g
RS	R + S	S3_g	S6_g + S7_g
Rs	R + s	S4_g	S6_g + S8_g

Table 9.: Scenario comparison within the CARP and the SP-CARP K, N, and S sets.

Secondly, it is interesting to compare the scenario where each household is free to choose the collection interval that suits best to a scenario in which all households are forced to have the same collection interval. This comparison can be performed directly by using the SP-CARP instances that contain the original flexible demand and the CARP graphs where all demand is normalized to bi-weekly service. There is a unique correspondence between the graphs for the two problems, which is indicated by the names. Furthermore, the capacities and compression factors of the vehicles for the two problems are matched such that vehicle 1 for the CARP matches vehicles 1-2, 1-4, and 1-8 for the SP-CARP, the only difference being that the number of vehicles used in the CARP (which is infinite for our instances) should be reduced to fit the availability from the SP-CARP. Thereby, 249 scenario comparisons are directly available across the two problems.

Thirdly, an interesting issue within waste collection is whether to sort and then collect or to collect and then sort. This can be analyzed using the No-Split MC-CARP instances, where sets A, B and D are directly comparable, collecting 2, 3, and 4 separated fractions, respectively. Similarly, instances R-A, R-B, and R-D are comparable. These instances provide comparison across 100 instances each. The vehicles assigned to each graph can be directly used by noting that vehicles 1-1-A, 1-1-B, and 1-1-D are identical, except from the way in which they are partitioned into compartments and the fact that compression factors are adjusted when fractions are split. This naming system is repeated throughout the vehicles.

Fourthly, for the C-CARP instances, it is quite highly relevant to compare the scenario with enforced coordination to the scenario with free choice of weekdays. Here, the latter reduces to solving a number of separate CARPs for the waste fractions using the vehicles assigned to that fraction. A comparison of these two scenarios can provide the cost of coordination which is interesting from a managerial perspective.

Another interesting issue for analysis is the effect of vehicle size change. This can be done directly for the CARP, the SP-CARP, and the No-Split MC-CARP by noting the scaling of the vehicles. For the CARP (and thereby for the SP-CARP), vehicle 1 (1-2, 1-4, and 1-8 for the SP-CARP) is the smallest, and vehicle 6 (6-2, 6-4, 6-8, and 6-24 for the SP-CARP) is the largest. For the No-Split MC-CARP, the vehicle files have been named e.g. 1-2-D, where the letter refers to the set for which it was created. The first number in the name is related to the relative size of the compartments of the vehicle and the compression, and no comparison should be made across them (unless one purposely wants to do so, of course). The second number in the name relates to the size of the vehicle (1 being the smallest) such that for instance vehicles 1-1-D, 1-2-D, 1-3-D, and 1-4-D are similar except from a scaling factor (not exactly, but close enough), providing total capacities of $21.8m^3$, $26.1m^3$, $31.3m^3$, and $40.0m^3$. Similarly for all other vehicles used for this problem.

Finally, a frequent issue of debate is that of separate or joint collection with single and multi-compartment vehicles, respectively. In all three MC-CARP versions, the E, R-E, F, and R-F graphs contain all 6 (or 7) waste fractions completely separated. Based on those graphs, any analysis of collection jointly or separately in any desired combination can be performed. Similarly, an analysis of joint collection of general waste and paper in the two-compartment vehicles versus separate collection of the two fractions versus collection of the two fractions in a mix can be performed directly on set A by assigning vehicles that fits the analysis.

6. Concluding Remarks

Five new variations of the CARP problem have been presented in this paper. Among these are three multi-commodity problems with multiple compartments in the vehicles: No-Split, Commodity-Split, and Multi-Day Commodity-Split. These three variations model a multi-commodity version of the CARP but include different degrees of complexity in order to model real-world waste collection problems in increasing detail. We also introduced a fourth multi-commodity version in which each commodity is to be collected by separate vehicles that have to coordinate collection. The final variation presented is a Semi Periodic CARP in which demand on edges is to be collected with a given interval and each edge may have demand with several different collection intervals simultaneously.

In addition, we have introduced large-scale test instances for each of these new CARP variations and for the original CARP. These instances originate from real-world data on waste collection in 5 areas in Denmark. The test instances have been created in a way that enables the decision maker to test various alternatives within waste collection. Different scenarios can be compared by solving at least two instances. Many of the graphs contain more than 1000 nodes and required edges. This implies that they are much larger than the classical CARP instances.

The focus in this paper has been on presenting the new problems and on providing large-scale instances for each of them based on real-life data. Therefore, an obvious next step would be to find solution methods that are able to solve these problems, both as regards the added complexity of the problem compared to the classical CARP and as regards the size. The size of the problems makes construction of meta heuristics and optimized districting approaches natural choices for future research. The paper thus provides a sound basis that can be used in future research that aims at closing the gap between academic research and the real-life challenges of large complex arc

routing problems.

Funding

The project is funded by the Danish Council for Independent Research Social Sciences project Transportation issues related to waste management. We are very grateful to our collaborators, Frederiksberg county (Frederiksberg county 2017), Odense Renovation (Odense Renovation A/S 2017), Reno Djurs (Reno Djurs 2017), and Renosyd (Renosyd 2017) for sharing their data and their knowledge with us and thereby making this project possible.

References

- Amponsah SK, Salhi S. 2004. The investigation of a class of capacitated arc routing problems: the collection of garbage in developing countries. *Waste Management*. 24(7):711 – 721.
- Aráoz J, Fernández E, Franquesa C. 2013. Grasp and path relinking for the clustered prize-collecting arc routing problem. *Journal of Heuristics*. 19(2):343–371.
- Bautista J, Fernández E, Pereira J. 2008. Solving an urban waste collection problem using ants heuristics. *Computers & Operations Research*. 35(9):3020 – 3033. Part Special Issue: Bio-inspired Methods in Combinatorial Optimization.
- Belenguer JM, Benavent E, Lacomme P, Prins C. 2006. Lower and upper bounds for the mixed capacitated arc routing problem. *Computers & Operations Research*. 33(12):3363 – 3383. Part Special Issue: Recent Algorithmic Advances for Arc Routing Problems.
- Benavent E, Campos V, Corberán Á. 1992. The Capacitated Arc Routing Problem: Lower Bounds. *Networks*. 22:669–690.
- Brandão J, Eglese RW. 2008. A deterministic tabu search algorithm for the capacitated arc routing problem. *Computers & Operations Research*. 35(4):1112–1126.
- Constantino M, Gouveia L, Mourão MC, Nunes AC. 2015. The mixed capacitated arc routing problem with non-overlapping routes. *European Journal of Operational Research*. 244(2):445 – 456.
- Corberán Á, Fernández E, Franquesa C, Sanchis JM. 2011. The windy clustered prize-collecting arc-routing problem. *Transportation Science*. 45(3):317–334.
- Corberán Á, Laporte G. 2015. Arc routing: problems, methods, and applications. Vol. 20. SIAM.
- Coutinho-Rodrigues J, Rodrigues N, Clímaco J. 1993. Solving an urban routing problem using heuristics: a successful case study. *Belgian Journal of Operations Research, Statistics and Computer Science*. 33(1):2.
- Del Pia A, Filippi C. 2006. A variable neighborhood descent algorithm for a real waste collection problem with mobile depots. *International Transactions in Operational Research*. 13(2):125–141.
- Dos Santos CG, Scarpin CT, Da Silva FP. 2016. Mathematical modeling for periodic capacitated arc routing problem. *International Journal of Engineering*. 8(02):8269.
- Econet AS. 2012. Idekatalog - Ordninger for dagrenovation. Only available in Danish.
- Eglese RW, Murdock H. 1991. Routing Road Sweepers in a Rural Area. *Journal of Operational Research Society*. 42(4):281–288.
- Frederiksberg county. 2017. Webpage; [<http://www.frederiksberg.dk/affald>]. Only available in Danish.
- Ghiani G, Guerriero F, Improta G, Musmanno R. 2005. Waste collection in southern italy: solution of a real-life arc routing problem. *International Transactions in Operational Research*. 12(2):135–144.

- Ghiani G, Improta G, Laporte G. 2001. The capacitated arc routing problem with intermediate facilities. *Networks*. 37(3):134–143.
- Ghiani G, Laganà D, Laporte G, Mari F. 2010. Ant colony optimization for the arc routing problem with intermediate facilities under capacity and length restrictions. *Journal of Heuristics*. 16(2):211–233.
- Golden BL, DeArmon JS, Baker EK. 1983. Computational Experiments with Algorithms for a Class of Routing Problems. *Computers and Operations Research*. 10(1):47–59.
- Golden BL, Wong RT. 1981. Capacitated Arc Routing Problems. *Networks*. 11:305–315.
- Gouveia L, Mourão MC, Pinto LS. 2010. Lower bounds for the mixed capacitated arc routing problem. *Computers & Operations Research*. 37(4):692 – 699.
- Kiushi M, Shinano Y, Hirabayashi R, Saruwatari Y. 1995. An Exact Algorithm for the Capacitated Arc Routing Problem using Parallel Branch and Bound Method. Abstract of the 1995 Spring National Conference of the Operational Research Society of Japan:28–29.
- KortCenter A/S. 2017. Webpage; [<http://www.kortcenter.dk>]. Only available in Danish.
- Li LYO, Eglese RW. 1996. An Interactive Algorithm for Vehicle Routeing for Winter-Gritting. *Journal of the Operational Research Society*. 47:pp. 217–228.
- McBride R. 1982. Controlling left and u-turns in the routing of refuse collection vehicles. *Computers & Operations Research*. 9(2):145 – 152.
- Mei Y, Tang K, Yao X. 2011. A Memetic Algorithm for Periodic Capacitated Arc Routing Problem. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*. 41(6):1654–1667.
- Mourão MC, Almeida MT. 2000. Lower-bounding and heuristic methods for a refuse collection vehicle routing problem1. *European Journal of Operational Research*. 121(2):420 – 434.
- Mourão MC, Amado L. 2005. Heuristic method for a mixed capacitated arc routing problem: A refuse collection application. *European Journal of Operational Research*. 160(1):139 – 153.
- Mourão MC, Nunes AC, Prins C. 2009. Heuristic methods for the sectoring arc routing problem. *European Journal of Operational Research*. 196(3):856 – 868.
- Odense Renovation A/S. 2017. Webpage; [<http://www.odenserenovation.dk/>]. Only available in Danish.
- Polacek M, Doerner KF, Hartl RF, Maniezzo V. 2008. A variable neighborhood search for the capacitated arc routing problem with intermediate facilities. *Journal of Heuristics*. 14(5):405–423.
- Reno Djurs I/S. 2017. Webpage; [<http://www.renodjurs.dk/>]. Only available in Danish.
- Renosyd. 2017. Webpage; [<http://www.renosyd.dk/>]. Only available in Danish.
- Santos L, Coutinho-Rodrigues J, Antunes CH. 2011. A web spatial decision support system for vehicle routing using google maps. *Decision Support Systems*. 51(1):1 – 9.
- Zhang Y, Mei Y, Tang K, Jiang K. 2016. Memetic algorithm with route decomposing for periodic capacitated arc routing problem. *Applied Soft Computing*. 52:1130 – 1142.