A simheuristic algorithm for time-dependent waste collection management with stochastic travel times

Aljoscha Gruler, Antoni Pérez-Navarro, Laura Calvet, and Angel A. Juan

Abstract

A major operational task in city logistics is related to waste collection. Due to large problem sizes and numerous constraints, the optimization of real-life waste collection problems on a daily basis requires the use of metaheuristic solving frameworks to generate near-optimal collection routes in low computation times. This paper presents a simheuristic algorithm for the time-dependent waste collection problem with stochastic travel times. By combining Monte Carlo simulation with a biased randomized iterated local search metaheuristic, time-varying and stochastic travel speeds between different network nodes are accounted for. The algorithm is tested using real instances in a medium-sized city in Spain.

MSC: 90B06, 68U20, 68T20, 90C59.

Keywords: Waste collection management, vehicle routing problem, stochastic optimization, sim-heuristics, biased randomization, case study.

1 Introduction

Due to its high operational costs and numerous related negative externalities such as air pollution, noise, and traffic congestion, waste management is among the most important public services (Strand, Syberfeldt and Geertsen, 2020). The complete process of collecting and disposing different types of garbage is a complex task shaped by various optimization problems related to facility location, clustering of service territories, and vehicle routing (Ghiani et al., 2014). Considering rising population numbers in urban areas around the world, especially waste collection processes need to be organized in an efficient manner in order to ensure a sustainable, cost-efficient, and...
citizen-friendly metropolitan garbage collection (Bing et al., 2016). The waste collection problem (WCP) is a rich extension of the well-known vehicle routing problem (VRP) with the aim of minimizing a certain objective function, e.g.: distances, travel times, CO₂ emissions, etc. (Kim, Kim and Sahoo, 2006). Problem inputs include a set of waste containers that hold a positive amount of waste, which has to be collected from a number of capacitated garbage collection vehicles located at a central depot. Moreover, the problem setting includes one or more landfills at which collected waste is disposed if a vehicle is full or before it returns to the central depot.

Given the practical nature of the WCP, realistic problem instances discussed in the literature typically include several hundred waste containers and several constraints related to maximum route travel times, driver lunch breaks, time windows, etc. (Benjamin and Beasley, 2010; Buhrkal, Larsen and Ropke, 2012). This imposes certain limits on the use of exact methods to solve this NP-hard problem, calling for the application of metaheuristic algorithms that are able to generate near-optimal solutions to large-scaled and realistic WCP settings in calculation times of only a few seconds or minutes. However, most metaheuristic solving methodologies still make simplifying assumptions about the nature of input variables. On the one hand, most routing optimization frameworks assume travel times between different network nodes to be static over time. Especially in the context of daily collection of waste, this is an unrealistic assumption due to the natural time dependency of edge traversing duration and vehicle velocities (Gendreau, Ghiani and Guerriero, 2015). On the other hand, a frequent drawback of many solving approaches is that they do not consider uncertainty in input variables. In the context of vehicle routing, information regarding travel times, demands, or customers themselves is typically not perfectly known in advance. Indeed, they are more likely to be of stochastic or even dynamic nature (Pillac et al., 2013; Ritzinger, Puchinger and Hartl, 2016).

Figure 1 illustrates the effects of time-dependent and stochastic travel speeds. Given the distance of traversing any edge in a routing problem, the travel duration to pass this edge can be calculated as the quotient of travel distance and the expected vehicle speed. In time-dependent routing scenarios, driving velocities vary according to different time periods within the route planning horizon. Apart from the expected travel speeds, realistic problem settings should also consider travel time variances due to different levels of planning uncertainty. The effects of different travel time assumptions are highlighted as optimistic and pessimistic vehicle speeds below, showing that variances in vehicle velocities can significantly impact the necessary time to visit a number of nodes, whereas the traveled distance is the same in all cases. This input uncertainty naturally occurs in most real-life routing problems, especially in metropolitan areas where actual travel times between different points are almost impossible to predict. A solution for the time-dependent VRP with time windows was already proposed by Figliozzi (2012), although the work lacks of real time implementations as well as alternative route constructions.

In general, one of the main issues related to routing problems applied in an urban context with uncertainty related to the transportation costs is how to define realistic instances (Tadei, Perboli and Perfetti, 2017). Usually algorithms are compared bet-
ween them by using a common database. Although this is very helpful to compare algorithms, it is mandatory to connect the algorithm with real data from users to apply them.

This paper presents a simheuristic approach (Juan et al., 2018) to solve the time-dependent WCP with stochastic travel times (TDWCPST). By integrating Monte Carlo simulation into a metaheuristic framework, both time dependencies and stochastic travel speeds can be accounted for. Our metaheuristic framework combines biased randomization techniques (Quintero-Araujo et al., 2017) with an iterated local search algorithm (ILS) by Lourenço, Martin and Stützle (2003). The inclusion of a simulation procedure during the optimization process leads to a couple of advantages over traditional metaheuristic solving approaches. Apart from the consideration of stochastic travel times, it allows for a closer statistical risk analysis of the obtained solutions. This enables the creation of additional decision-making dimensions related to route robustness in uncertainty scenarios, e.g.: standard deviations or different quartiles obtained during the simulation phase. The implementation and performance of the solving methodology is tested on a large-scale case study. This case study refers to the waste collection process in the medium-sized city of Sabadell, which is located within the autonomous region of Catalonia, in northern Spain. It is important to note that real data from the waste collector department of the city is transformed to create real instances where to apply the algorithm.

Thus, the contributions of this work are threefold: (i) motivated by a real-life case, a rich TDWCPST is proposed; (ii) a large real-life data set, with several realistic routing constraints, is used to show the applicability of the proposed optimization procedure; and (iii) the potential of the simheuristic approach is illustrated in a range of computational experiments, hence yielding various managerial insights.

\textbf{Figure 1: The effect of stochastic travel duration due to time-varying vehicle speeds in time-dependent routing scenarios.}
The paper is structured as follows: relevant literature on metaheuristic approaches for time-dependent routing problems and waste collection is reviewed in Section 2; the TDWCPST and the real-life problem setting are detailed in Section 3; Section 4 outlines our simheuristic solving framework; Section 5 describes different computational experiments and analyses obtained results; finally, Section 6 concludes this work and discusses possible future research directions.

2 Literature Review

This section reviews recent literature regarding metaheuristic solving frameworks for time-dependent VRPs and the WCP. For a more detailed overview on previous research regarding time-dependent routing problems the reader is referred to the work of Gen-dreau et al. (2015). A more extensive literature review on operational challenges and optimization methodologies in waste management is provided by Beliën, De Boeck and Van Ackere (2014) and Han and Ponce-Cueto (2015).

2.1 Metaheuristic solving methodologies for time-dependent routing problems

In the field of vehicle routing optimization, time dependency was not considered up to the early 2000s apart from a few exceptions. Malandraki and Daskin (1992) formulated travel times as a step function of the time of the day. This approach has the major drawback that the no-passing, first-in-first-out (FIFO) property is not guaranteed. Thus, a vehicle leaving node $i$ might arrive later at node $j$ than a vehicle leaving node $i$ at a posterior starting time due to varying travel times. This drawback in the travel time function was improved by Hill and Benton (1992), who developed the first travel time model based on time-varying vehicle speeds, which implies the FIFO characteristic. Later, Ichoua, Gendreau and Potvin (2003) used an improved version of this vehicle speed model in combination with a parallel tabu search heuristic to show the benefits of time-dependent vehicle routing compared to its static counterpart. The impact of time-dependent travel times to avoid traffic congestion was also studied by Kok, Hans and Schutten (2012), who showed that late arrivals at customers and extra duty times through traffic jams can be significantly reduced through smart congestion-avoidance strategies.

An iterated local search algorithm for the time-dependent VRP with time windows (TDVRPTW) was presented by Hashimoto, Yagiura and Ibaraki (2008). Computational experiments include a variety of problem instances with up to 1,000 nodes. The TDVRPTW was also addressed in the works of Balseiro, Loiseau and Ramonet (2011) and Harwood, Mumford and Eglese (2013). The former developed an ant colony system hybridized with insertion heuristics which is tested on problem instances with up to 100 clients. The latter established quick estimates of time-dependent travel times for
the traveling salesman problem. Their results show that their estimations can lead to significant reductions in computation time. The TDVRP with simultaneous pickup and deliveries was addressed by Zhang, Chaovalitwongse and Zhang (2014) through an integrated ant colony and tabu search approach. A total of 100 customers were considered in their work. During the last decade, much attention has also been paid to the environmental effects of routing, in the context of the so called pollution routing problem. Kuo (2010) developed a simulated annealing algorithm for establishing emission minimizing vehicle routes while taking into account varying edge traversing times. Computational results are provided using benchmark instances with up to 100 customers. The trade-off between travel times and CO$_2$ emissions in time-dependent VRPs was analyzed by Jabali, Van Woensel and de Kok (2012). The time-dependent pollution routing problem was also analyzed in the work of Franceschetti et al. (2013). The authors proposed an integer linear programming formulation for cases without any traffic congestion. Environmental considerations are also included in the work of Soysal, Bloemhof-Ruwaard and Bektas (2015), who addressed the time-dependent two-echelon VRP through a comprehensive mixed integer linear programming (MILP) formulation.

All previously cited works focused on the deterministic version of the TDVRP. For stochastic problem settings the literature is more scarce. Lecluyse, Van Woensel and Peremans (2009) developed a tabu search metaheuristic for the TDVRP with stochastic travel times. Nahum and Hadas (2009) developed an extended version of the well-known savings algorithm to address the stochastic TDVRP. Tas et al. (2014) proposed a tabu search and adaptive large neighbourhood search metaheuristic for the TDVRP with soft time windows and stochastic travel times.

### 2.2 Metaheuristic frameworks in the optimization of waste collection

Different metaheuristic approaches have been presented in the solution of various WCPs and their extensions. Even though many works include a case study to show the real-life potentials of their frameworks, to the best of our knowledge, time dependency in the WCP has not yet been considered in the literature.

Baptista, Oliveira and Zúquete (2002) elaborated an extension of the Christofides and Beasley heuristic for the multi-period WCP modeled as a periodic VRP (PVRP) to combine vehicle scheduling over multiple time periods with route planning. The authors used their approach to improve municipal waste collection in the Portuguese city of Almeda. Also addressing a multi-period WCP, Teixeira, Antunes and de Sousa (2004) developed a cluster-first route-second heuristic to schedule and plan waste collection routes for different waste types in a case study in Portugal with over 1600 collection sites. Nuortio et al. (2006) presented a guided variable thresholding metaheuristic to solve a multi-period WCP with several thousand collection points in Eastern Finland. Hemmelmayr et al. (2013) addressed the PVRP with different waste types and up to 288 containers, which they solved with a variable neighbourhood search metaheuristic. They consider the landfills as intermediate facilities, which are inserted in pre-constructed
A simheuristic algorithm for time-dependent waste collection management... routes using dynamic programming. In the same work, the authors also discussed the single period WCP with multiple depots, in which the landfills serve as vehicle depots and disposal sites at the same time. Ramos, Gomes and Barbosa-Póvoa (2014) extended the typical objective of minimizing routing costs in order to include environmental concerns, considering multiple waste types and numerous vehicle depots in a case study in Portugal.

Only focusing on waste collection routing, Kim et al. (2006) developed an extension of Solomon’s insertion algorithm to optimize routes of a North American waste management service provider, considering a capacitated vehicle fleet, time windows, and driver lunch breaks. The authors reported reduced routing distances of up to 10%. Furthermore, a benchmark set of 10 realistic instances based on the original case study ranging from 102 TO 2,100 nodes is provided. Using the same benchmark set, Benjamin and Beasley (2010) combined tabu search with a variable neighbourhood search metaheuristic. By exchanging containers and landfills within and between routes, the solution search space is systematically increased. Likewise, Buhrkal et al. (2012) proposed an adaptive large neighbourhood search metaheuristic. Based on an initial solution, their approach applies a range of destroy-and-repair methods to examine several solution neighbourhoods. It is called adaptive since the choice of methods depends on the solution quality obtained during the construction of earlier solutions. Moreover, an acceptance criterion for new solutions based on simulated annealing is included. Likewise, Markov, Varone and Bierlaire (2016) presented a multiple neighbourhood search heuristic for a real-world application of the waste collection VRP with intermediate facilities. The authors consider a heterogeneous vehicle fleet and flexible depot destinations in their approach. Gruler et al. (2017a) developed a metaheuristic algorithm to assess the potentials of horizontal collaboration in urban waste collection.

Concerning the WCP under input uncertainty, the literature is more scarce with most works focusing on stochasticity concerning expected waste levels. Ant colony optimization and a hybrid approach based on a genetic algorithm and tabu search for a case study with 50 containers in Malaysia is presented in Ismail and Irhamah (2008) and Ismail and Loh (2009). After planning aprioristic routes, waste levels are simulated according to a discrete probability distribution. Routes undergo a recourse action (i.e., an additional disposal trip) whenever actual demand exceeds the planned collection amount. Nolz, Absi and Feillet (2014) formulated a collector-managed inventory routing problem for a case study on the collection of infectious waste. By using real information obtained through radio frequency identification, their adaptive large neighbourhood search algorithm is able to consider stochastic waste collection levels. Alshraideh and Abu Qdais (2017) combined a multi-period WCP with time windows and stochastic demands in a real case study of medical waste collection from 19 hospitals in Northern Jordan. They used a genetic algorithm and a probability constraint regarding a pre-defined service level to solve the problem. Also, Gruler et al. (2017b) presented a variable neighbourhood search based simulation-optimization approach for the WCP with stochastic demands. Although metaheuristics are becoming the predominant methodology in solving
WCP under rich and realistic scenarios (Hannan et al., 2018; Asefi et al., 2019), other approaches such as mixed-integer programming are also being employed by some experts (Mohsenizadeh, Tural and Kentel, 2020).

3 Problem Description

This section outlines the real-life case study of collecting waste under different routing constraints and travel time assumptions in the city of Sabadell. Furthermore, the time-dependent WCP with stochastic travel times (TDWCPST) and the applied travel speed model for different time periods is discussed in more detail.

3.1 The waste collection problem in Sabadell

Sabadell is a medium-sized city of roughly 200,000 inhabitants located within the autonomous Spanish region of Catalonia. Collection vehicles are located at a central depot and collected garbage is disposed in a single landfill. Expected waste levels in each container, average service times at each node, and the average vehicle travel speeds during different time periods are known. The problem settings consists of a total of 921 paper waste containers which are currently visited on 9 different routes. The locations of the vehicle depot, the landfill, waste containers, and the original route assignment can be seen in Figure 2 (the central depot and the landfill are marked by the square symbols).

According to the managers, in a scenario with dynamic travel times as the one being considered, the total time required to complete the waste collection process is the main key performance indicator. On the one hand, the operational times directly affect the operational costs associated with the waste collection process in terms of wages and vehicle usage costs. On the other hand, an important routing constraint is that collection routes need to be completed between 9 a.m. and 4 p.m., as these are the opening hours of the central depot at which the collection vehicles are stationed. Moreover, different time periods within the daily planning horizon can be identified regarding expected traffic speeds:

- Heavy traffic on all streets is expected during the rush hour from 9 a.m. to 10 a.m. and from 1 p.m. to 2 p.m.
- Traffic jams are expected in streets close to primary schools in the time periods of 9 a.m. to 10 a.m., 12 p.m. to 1 p.m., and 3 p.m. to 4 p.m.

Especially the latter observation is of importance in the planning of waste collection routes. Containers in the affected streets should not be visited within the depicted time period. Due to parents picking up their children from primary schools, streets within a certain distance radius of the school building should be avoided in the given period if
possible. According to the experience of the decision-taker, a radius of 500 m around primary schools is considered. Apart from delays in the collection process, visiting these streets during the most busy hours affects many citizens and can even be dangerous due to children exiting the primary school facilities. The influenced streets in the city centre of Sabadell for which the additional constraints apply are highlighted in Figure 3.

3.2 A time dependent travel speed model for the WCP

The TDWCPST can be described on a graph $G = (V, E)$:
Node set $V = V^d \cup V^f \cup V^c$ includes:

(i) A central depot $V^d = \{0\}$ at which a homogeneous fleet of waste collection vehicles, each of them with capacity $C$, is located.

(ii) A set $V^f = \{1, 2, \ldots, m\}$ of $m$ landfills at which collected waste must be disposed if vehicle capacities are reached and before a vehicle returns to the central depot (making more than one landfill trip per route possible if no other route constraints are violated).
(iii) A set of $n$ waste containers $V^c = \{m+1, \ldots, m+n\}$ with associated waste levels $q_i > 0$ ($\forall i \in V^c$). Service times for emptying any container and for disposing collected garbage at any landfill are defined as $s_i > 0$ ($\forall i \in V \setminus V^d$).

- Edge set $E = \{(i, j) / i, j \in V, i \neq j\}$ describes all edges connecting any two nodes.
- Travel distances $d_{ij} \geq 0$ between any two nodes in $V$ are known.
- Additional routing constraints include a maximum amount of waste to be collected during each route and the maximum route duration defined by the opening and closing times at the central depot.

Our travel speed model for time varying vehicle velocities is based on the discussions of Ichoua et al. (2003). The planning horizon (defined by the depot opening hours) is divided into $p$ time periods $T_1, T_2, \ldots, T_p$. Travel durations $tt(T)$ to cross any edge in $e \in E$ can be calculated as the quotient of travel distances and vehicle speeds $v_T$ ($T \in \{T_1, T_2, \ldots, T_p\}$), such that $tt(T) = d_e / v_T$. In the specific case of waste collection in Sabadell, different travel speeds can be defined for different edges, e.g., due to rush hour traffic or other events such as opening or closing hours of schools. For this reason, edge set $E$ is partitioned into $S$ subsets with $E_s$ ($s = 1, 2, \ldots, S$). Thus, travel speeds can be formulated as $tt_s(T)$ to show the travel speed of any edge of the edge subset $E_s$ during time period $T$. This step-wise travel speed model along different times of the planning horizon is a natural way of estimating travel duration of different edges in real-world conditions. Furthermore, it implies the satisfaction of the FIFO property.

4 Solving Framework

The different stages of the proposed simheuristic solving methodology for the TD-WCPST are summarized in Figure 4. By integrating simulation into a biased-randomized iterated local search (BR-ILS) algorithm, a set of promising stochastic solutions are constructed. These solutions are then refined in a more intensive simulation procedure. Finally, the defined set of solutions undergoes a more detailed risk analysis according to different criteria. All steps are outlined in more detail in the following subsections.

4.1 Constructing an initial time-dependent WCP solution

Our approach starts by constructing a feasible initial solution with an enhanced framework of the well-known savings heuristic for routing problems (Clarke and Wright, 1964). In the original procedure, the savings $s_{ij}$ of including any edge $e$ connecting two customers $i$ and $j$ in a constructed solution are calculated as $s_{ij} = s_{ji} = c_{i0} + s_{0j} - c_{ij}$. However, this assumption does not hold in the special case of waste collection, as the round trip costs between the central depot and any waste container are asymmetric due to the additional landfill visit at the end of any completed route. Thus, the route travel
direction influences the savings values assigned to each edge. In order to account for the necessary landfill visit in every route, the expected savings of each edge are calculated as average values of completing a route in both directions, such that \( E[s_{ij}] = (s_{ij} + s_{ji}) / 2 \). After each merge, this initial estimate is updated to account for the real travel times depending on the hour of the day.

![Simheuristic solving methodology.](image)

Apart from this algorithm adaption to the problem setting, we enhance the greedy edge selection process of the savings procedure through a probabilistic construction behaviour based on biased randomization techniques (Ferone et al., 2019). As highlighted in Algorithm 1, a candidate set of edges is ranked according to their respective savings value. In the following, a feasible waste collection route is created by iteratively adding solution elements from the eligible edges. Selection probabilities follow a geometric distribution defined through parameter \( \alpha (0 < \alpha < 1) \), which depicts the probability of the most promising solution element to be chosen. This process is similar to...
the GRASP procedure discussed in the work of Resende and Ribeiro (2010). However, while GRASP is based on a restricted candidate list and a uniform selection probability, selection probabilities are inclined to more promising solution elements—which are all potentially eligible at each solution construction step—in this biased randomization approach. In a biased-randomized algorithm, the choice of the skewed probability distribution has an impact on the quality of the final solution. As discussed in Grasas et al. (2017), the geometric and the decreasing triangular probability distributions have been successfully used in previous work, but other probability distributions (either theoretical or empirical) are possible as well.

Algorithm 1: Biased randomization to create an initial TDWCP solution

<table>
<thead>
<tr>
<th>Input: Skewed probability distribution $f$, parameter $\alpha$, edge set $E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 $sol \leftarrow \emptyset$</td>
</tr>
<tr>
<td>2 initialize candidate set: $CL \leftarrow E$</td>
</tr>
<tr>
<td>3 sort $CL$ according to savings value</td>
</tr>
<tr>
<td>4 while solution $sol$ is not complete do</td>
</tr>
<tr>
<td>5 Randomly select $pos \in CL$ according to distribution function $f(\alpha)$</td>
</tr>
<tr>
<td>6 $sol \leftarrow sol \cup pos$</td>
</tr>
<tr>
<td>7 $CL \leftarrow CL \setminus pos$</td>
</tr>
<tr>
<td>8 sort $CL$</td>
</tr>
<tr>
<td>end</td>
</tr>
<tr>
<td>9 return TDWCP solution $sol$</td>
</tr>
</tbody>
</table>

4.2 A simheuristic framework for the time-dependent WCP with stochastic travel times

Our simheuristic procedure to solve the TDWCPST is outlined in Algorithm 2. Once an initial solution is constructed and set as the current incumbent $baseSol$ and $bestSol$ solutions, an iterated local search algorithm is started (Lourenço et al., 2003). During a predefined stopping criterion, new TDWCP solution neighbourhoods are created by perturbing the current $baseSol$. Each perturbated solution $newSol$ then undergoes a local search phase to find the local minimum within the current solution structure. As perturbation operator a double-bridge move is applied. Hereby, a solution is partitioned into four pieces of random size, which are subsequently joined in an arbitrary order. As local search movement, a $2$-opt operator is employed (Muyldermans et al., 2005).

Up to this point, deterministic (expected) travel duration between difference network nodes are considered. In order to account for uncertainty in input variables, Monte Carlo simulation is applied to any promising solution found in the metaheuristic search. A TDWCP solution $newSol$ is deemed promising if its deterministic travel duration outperform those of the currently incumbent $baseSol$ or if if a simulated annealing-like acceptance criterion is met. The travel duration between all edges of a promising solution are simulated from a log-normal probability distribution during $nSim$ simulation runs. At this stage any other probability distribution could be applied, but the log-normal one is a “natural” choice to model non-negative random variables, such as travel times...
in routing problems or times-to-failure in reliability studies (Faulin et al., 2008). During each simulation iteration, expected travel duration $tt_{t}$ between any two points are defined as distribution mean of the probability function. Variance factor $k$ defines travel duration variance levels. With $E[tt_{t}] = tt_{t}$ and $Var[tt_{t}] = k \cdot tt_{t}$, the location parameter $\mu_{i}$ and scale parameter $\sigma_{i}$ defined for the probability function can be formulated as:

$$
\mu_{i} = \ln(E[tt_{t}]) - \frac{1}{2} \ln \left( 1 + \frac{Var[tt_{t}]}{E[tt_{t}]^2} \right)
$$

$$
\sigma_{i} = \left\lfloor \sqrt{\ln \left( 1 + \frac{Var[tt_{t}]}{E[tt_{t}]^2} \right)} \right\rfloor
$$

As a result of time varying travel speeds, the variability in solution waste collection durations estimated after each simulation run can be expected to increase with higher variance levels. In particular, waste collection close to primary school locations is penalized by significantly reduced travel speeds during predefined time periods. After the simulation phase, the stochastic travel durations of newSol are defined as the average of all simulation results.

If the stochastic costs of the considered solution outperform the estimated stochastic travel durations of the incumbent baseSol and/or bestSol, they are updated respectively. Moreover, each solution that is defined as incumbent baseSol during any stage of the simheuristic procedure is included in a TDWCPST solution set eliteSols. After the algorithm stopping criterion is reached, solutions included in this exclusive set of elite solutions undergo a more intensive simulation phase defined by a higher number of simulation runs. This allows a more accurate estimation of the best found WCP solutions in stochastic travel time scenarios.

The described combination of simulation with metaheuristics leads to several advantages over deterministically focused optimization approaches. Firstly, the search phase is driven by the stochastic solution estimates obtained during the simulation (i.e., the baseSol is updated according to the cost estimates provided by the simulation component). Secondly, TDWCPST solutions can be realistically evaluated under different uncertainty scenarios. Finally, the simheuristic methodology allows the evaluation of different solutions according to additional criteria instead of simply focusing on the defined objective function. Due to the stochasticity in real-life travel duration, the completion of waste collection plans is likely to vary with respect to the predicted driving times. For this reason, decision-makers need a more insightful decision support than simply focusing on the minimization of expected travel times. Thus, we implement a final risk analysis for the elite solutions in our simheuristic procedure. At this stage, additional dimensions related to the robustness of a considered solution, such as the standard deviation or the quartiles, are computed.
Algorithm 2: A simheuristic for the TDWCPST

Input: \( f, E, \alpha, n_{\text{Sim, short}}, n_{\text{Sim, long}}, k \)

1. \( \text{nodes} \leftarrow \text{getNodes}(E) \)
2. \( \text{costMatrix} \leftarrow \text{getCostMatrix}(E) \)
3. \( \text{initSol} \leftarrow \text{generateBRSolution}(f, \alpha, E) \)  // Biased-Randomized Algorithm
4. baseSol \leftarrow \text{initSol}
5. stochDuration(baseSol) \leftarrow \text{infinite}
6. bestSol \leftarrow \text{baseSol}
7. elitSols \leftarrow \emptyset
8. while stopping criterion not reached do
9.     newSol \leftarrow \text{perturbate(baseSol, costMatrix)}  // perturbation stage
10.    newSol \leftarrow \text{localSearch(newSol, costMatrix)}  // local search stage
11.    delta \leftarrow \text{detDuration(baseSol)} - \text{detDuration(newSol)}
12.    if delta \geq 0 then
13.        credit \leftarrow \text{delta}
14.        stochDuration(newSol) \leftarrow \text{simulation(newSol, n_{\text{Sim, short}}, k)}
15.        if stochDuration(newSol) \leq stochDuration(baseSol) then
16.            includeInEliteSolutionSet(newSol)
17.            baseSol \leftarrow newSol  // simulation driven baseSol
18.        if stochDuration(newSol) < stochDuration(bestSol) then
19.            bestSol \leftarrow newSol
20.    end
21.  end
22.  else if \text{delta} \leq 0 then
23.      credit \leftarrow 0
24.      stochDuration(newSol) \leftarrow \text{simulation(newSol, n_{\text{Sim, short}}, k)}
25.      baseSol \leftarrow newSol
26.  end
27.  for eliteSol \in \text{elitSols} do
28.      stochDuration(eliteSol) \leftarrow \text{simulation(eliteSol, n_{\text{Sim, long}}, k)}
29.  end
30. return bestSol

4.3 Creating a real-life distance matrix

In this subsection we show the process followed to generate a real-life distance matrix. The data was obtained through a collaboration agreement between the Internet Computing and Systems Optimization (ICSO@IN3) research group and the company SMATSA, which is responsible for the collection of waste in the inner-city area of Sabadell. The problem dealt in the present paper involves the waste disposal vehicle routing in order to design efficient routes between 886 paper waste containers. A single depot and landfill are considered. Locations are given as Longitude/Latitude (Long/Lat) and postal addresses are also available. The goal is to create a real-life distance matrix from these data.

In the creation of this distance matrix only open software has been used. In particular, we have used: QGIS 2.18 (https://www.qgis.org) as a geographic information system (GIS); PostGIS (https://postgis.net), which is the geographic extension of the database PostgreSQL (https://www.postgresql.org); and pgRouting (https://pgrouting.org) to obtain the distances between pairs of locations. In addition, Open Street Map or OSM (https://www.openstreetmap.org) has been em-
ployed as a base map. The first step is to download the base map for the zone of Sabadell from OSM. This downloaded file is then processed with osm2po (https://osm2po.de) to transform it into a routable file. One of the outputs of the program is an SQL file that can be executed in PostgreSQL with the PostGIS extension. The output is a table that can be visualized in QGIS with the Add PostGIS table function.

At this step the map is available in a GIS and, although the aspect is visually correct, it does not have topology, which is required to connect nodes and to obtain real distances. The topology can be created with the pgrouting query pgr-createTopology, which can be run from QGIS thanks to the database manager plug in. Once the map has a topology, pgRouting can be run to obtain the shortest path between two given points. Figure 5 shows the uploaded map with the route between two points.

![Figure 5: Route between two points obtained with pgRouting in QGIS after uploading the base map and having introduced the topology.](image)

After this process, we had the map of the working area in a GIS and prepared to obtain the distance between any two points. The next step is uploading the location of the 886 containers, the depot and the landfill. Although the location is quite precise, there
may be some errors and it is important to verify out-layer nodes. Usually these points can be visually located in the map and corrected using the postal address (Figure 6).

On the other hand, the position of the nodes does not correspond to the street lines of the map in QGIS, since the streets have a width and they are represented as a line axe. Therefore, every point has to be linked to its corresponding street axis. It has been done with the \textit{NNJoin QGIS} plug in, which obtains the nearest neighbour between every single point and the start/end node of street axes. Figure 6 shows the nodes within the network and the imported nodes. Thus, it has been possible to obtain the closest point from the axes to every single container, which generates a new list of points, but this time, within the routable network. This introduces a little error in the position of the containers if they are not in the end node of the street axe. A more precise solution would have been to locate the nearest point of the axe and split the axe at that point. However, since containers are usually located at the corner of the streets, and according to the managers’ opinion, the error introduced can be considered as a non-relevant one for the purposes of this study.
Finally, we can obtain the distance matrix using PostgreSQL (with the extensions PostGIS and pgRouting). The distance in km is set as cost, and a distance matrix of the first 10 nodes is established. The name of the output file is sabadel.2po.4pgr.

5 Computational Experiments and Analysis of Results

The proposed simheuristic solving framework is applied to the real-life waste collection problem setting described in section 3. The algorithm is implemented as a Java application and tests are run on a personal computer with 4GB RAM and an Intel Pentium® processor with 2.16GHz. The necessary algorithm parameters to complete the described tests are specified as follows. These parameters have been obtained after a quick calibration based on the methodology proposed by Calvet et al. (2016):

- Skewed probability distribution \( f \): geometric with parameter \( \alpha = 0.3 \).
- \( nSim_{\text{short}} \): 100.
- \( nSim_{\text{long}} \): 1000.
- BR-ILS stopping criterion per instance: 30 seconds.

According to the observations of the decision maker, the average vehicle speed in normal traffic conditions is 25 km/h. The travel speed is divided by 5 and 25 during heavy traffic and traffic jams, respectively. Average vehicle service times at each container are set to 90 seconds, while 45 minutes are necessary to empty a vehicle at the landfill. Stochastic travel times are generated with three different variance factors, \( k = 1, 2, 5, 10 \), representing different (low / medium / high) uncertainty levels. All variance scenarios are represented in Figure 7, showing the travel times of edge subset \( s \) during time period \( T \) with an expected traversing time of \( E[tt_{st}] = 25 \) time units. The shaded area under each curve represents 95% of the simulated values. In the low-variance scenario \( (k = 1) \), 95% of actual driving times fall between 16.64 and 36.15 time units with a high density around the expected value. As the variance level is increased, the maximum density of the simulated times for all edges decreases and a higher variability can be observed. Note that the overall driving duration of a solution will increase as the expected travel time uncertainty increases. Moreover, the special case of \( k = 0 \) is equivalent to the deterministic routing case.

5.1 Experimental results

In order to evaluate the performance of our simheuristic algorithm, its results are compared to the nine waste collection routes currently completed on a daily basis in Sabadell. The comparison of the current routes (i.e., those built by SMATSA) and the best found solution of the BR-ILS algorithm in different variance scenarios is listed in Tables 1
(deterministic case and low variance) and 2 (medium and high variance). A total of 10 independent executions were run (each one using a different seed for the random number generator), and the best-found solution was returned by the algorithm. Each route holds between 79 and 122 waste containers to be emptied. To allow a fair comparison, the current order of visiting waste containers is evaluated in accordance with the necessary algorithm parameters described before. Since the distance matrix has been created using real-life distances, it is possible to make the comparison in order to know if the cost function is actually improved.

![Figure 7: Log-normal distribution of different variance levels around an expected travel time of 25 time units.](image)

It is important to note that comparing the nine routes separately, instead of designing new routes, allows us to compare our algorithm with the current situation in a real problem. The comparison is performed in terms of total time employed in completing the collection process, since this is the main key performance indicator for the managers.

In all travel duration variance scenarios, the BR-ILS is able to significantly outperform the current waste collection routes (by over 12% on average). Moreover, the solution travel duration in different uncertainty scenarios provided by our metaheuristic show that estimated travel duration increase with higher variance levels. The best results with the simheuristic BR-ILS algorithm are obtained when considering all 921 waste containers in a “global” waste collection instance. In this case, new route-to-containers assignments are established instead of solely focusing on reordering pre-established waste collection routes. For example, in the deterministic routing case, and
using a running time of 120 seconds, the global solution yields a overall driving duration of 2,820.5 minutes, with only 8 necessary garbage collection routes, thus saving one route to the company. Regarding the solution to the stochastic version of the problem, the proposed simheuristic has been run for a maximum time of 5 minutes before returning the best-found solution. This maximum computational time was suggested by the managers, who have to plan the collection routes every morning.

Table 1: Driving duration in minutes of current routes compared to our best found solution (deterministic case and low variance scenario).

<table>
<thead>
<tr>
<th>Route</th>
<th>k = 0</th>
<th>Diff (%)</th>
<th>k = 1</th>
<th>Diff (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Current</td>
<td>Our Best</td>
<td>(</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>392.5</td>
<td>357.3</td>
<td>−9.0</td>
<td>391.3</td>
</tr>
<tr>
<td>2</td>
<td>379.9</td>
<td>265.0</td>
<td>−30.2</td>
<td>381.3</td>
</tr>
<tr>
<td>3</td>
<td>470.8</td>
<td>345.5</td>
<td>−26.6</td>
<td>455.5</td>
</tr>
<tr>
<td>4</td>
<td>386.4</td>
<td>342.3</td>
<td>−11.4</td>
<td>384.2</td>
</tr>
<tr>
<td>5</td>
<td>374.3</td>
<td>335.1</td>
<td>−10.5</td>
<td>387.9</td>
</tr>
<tr>
<td>6</td>
<td>396.2</td>
<td>371.4</td>
<td>−6.2</td>
<td>397.1</td>
</tr>
<tr>
<td>7</td>
<td>372.8</td>
<td>340.4</td>
<td>−8.7</td>
<td>364.2</td>
</tr>
<tr>
<td>8</td>
<td>393.9</td>
<td>326.7</td>
<td>−17.0</td>
<td>399.8</td>
</tr>
<tr>
<td>9</td>
<td>407.9</td>
<td>323.0</td>
<td>−20.8</td>
<td>411.1</td>
</tr>
<tr>
<td>Total</td>
<td>3,574.5</td>
<td>3,006.7</td>
<td>−15.9</td>
<td>3,572.3</td>
</tr>
</tbody>
</table>

Table 2: Driving duration in minutes of current routes compared to our best found solution (medium and high variance scenario).

<table>
<thead>
<tr>
<th>Route</th>
<th>k = 2.5</th>
<th>Diff (%)</th>
<th>k = 10</th>
<th>Diff (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Current</td>
<td>Our Best</td>
<td>(</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>391.6</td>
<td>363.0</td>
<td>−7.3</td>
<td>392.7</td>
</tr>
<tr>
<td>2</td>
<td>381.4</td>
<td>281.3</td>
<td>−26.2</td>
<td>388.1</td>
</tr>
<tr>
<td>3</td>
<td>457.9</td>
<td>349.8</td>
<td>−23.6</td>
<td>460.2</td>
</tr>
<tr>
<td>4</td>
<td>383.9</td>
<td>342.6</td>
<td>−10.7</td>
<td>384.2</td>
</tr>
<tr>
<td>5</td>
<td>389.7</td>
<td>342.7</td>
<td>−12.0</td>
<td>395.4</td>
</tr>
<tr>
<td>6</td>
<td>397.2</td>
<td>386.9</td>
<td>−2.6</td>
<td>397.3</td>
</tr>
<tr>
<td>7</td>
<td>362.3</td>
<td>337.7</td>
<td>−6.8</td>
<td>360.4</td>
</tr>
<tr>
<td>8</td>
<td>398.8</td>
<td>344.0</td>
<td>−13.8</td>
<td>398.9</td>
</tr>
<tr>
<td>9</td>
<td>410.6</td>
<td>329.4</td>
<td>−19.8</td>
<td>414.0</td>
</tr>
<tr>
<td>Total</td>
<td>3,573.3</td>
<td>3,077.3</td>
<td>−13.9</td>
<td>3,591.1</td>
</tr>
</tbody>
</table>

5.2 Risk analysis of different TDWCPST solutions

For each waste collection plan (i.e., for each solution to the TDWCPSD), our simheuristic algorithm not only generates information about its expected travel duration and ex-
pected driving distance, but it can also provide the plan’s probabilistic profile (including risk and reliability analyses). Thus, statistical values such as the standard deviation of travel duration, the median, or the third quartile can be obtained during the simulation runs without increasing the computing effort.

Table 3 shows different attributes of three elite solutions of the global TDWCPSD with all garbage containers in a high variance scenario ($k = 10$). The deterministic and stochastic travel duration, driving distance, standard deviation, median, third quartile, and the number of waste collection routes of each TDWCPST solution are provided. As highlighted in the radar chart shown in Figure 8, each solution outperforms the others in a different decision-making dimension. While solution B is the most promising solution regarding deterministic travel duration, solution A shows the best results in terms of expected travel times and overall travel distance. However, the standard deviation of travel duration obtained during the long simulation run (which can be seen as a reliability indicator of a given solution) is the lowest for solution C. This solution behaviour is also observed in the multiple boxplot shown in Figure 9. It can be clearly seen that the most promising deterministic solution B yields the highest travel duration variance, suggesting a low reliability of the constructed waste collection routes. Likewise, the median and third quartile could be considered in a closer risk analysis according to the preferences of the waste collection route planner. Since this work is addressed to a real situation scenario, it is important for the planner this degree of freedom that allows to find different solutions. In this experiment we offered three solutions to the planner or decision maker. In other settings, the specific number could be adjusted taking into account the magnitude of the differences among solutions and the preferences of the planner.

Figure 8: Ranking of TDWCPST solutions according to different quality dimensions.
Table 3: Analysis of different TDWCPST solutions (high variance scenario).

<table>
<thead>
<tr>
<th>Solution</th>
<th>Det. Duration (min)</th>
<th>Stoch. Duration (min)</th>
<th>Distance (km)</th>
<th>Stand. Dev.</th>
<th>Median</th>
<th>Third Quartile</th>
<th># Routes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2,879.61</td>
<td>2,936.05</td>
<td>264.58</td>
<td>31.98</td>
<td>2,932</td>
<td>2,956</td>
<td>8</td>
</tr>
<tr>
<td>B</td>
<td>2,827.49</td>
<td>2,938.67</td>
<td>278.45</td>
<td>64.59</td>
<td>2,930</td>
<td>2,969</td>
<td>9</td>
</tr>
<tr>
<td>C</td>
<td>2,897.58</td>
<td>2,952.18</td>
<td>280.27</td>
<td>26.46</td>
<td>2,951</td>
<td>2,967</td>
<td>8</td>
</tr>
</tbody>
</table>

Figure 9: Comparison of simulation results of different TDWCPST solutions.

6 Conclusions

This work presents a simheuristic algorithm for the time-dependent waste collection problem with stochastic travel times to improve the real-life case of the waste collection process of several hundred waste containers in the Spanish city of Sabadell. The algorithm works by integrating simulation into a metaheuristic framework, which is based on a biased-randomized iterated local search. Uncertainty in travel duration between different nodes in the city logistics network is considered as well.

The work also shows the process followed to obtain the real-life distances. Working with real-life distances allows the comparison of the algorithm results with the real routes that are used in Sabadell nowadays. Results suggest significant travel duration reductions in different variance scenarios. Furthermore, a risk analysis of obtained solutions along different dimensions such as the standard deviation of travel duration is performed. The results underline the importance of risk aware route planning in the process of waste collection.
The research completed in this paper can be extended in several directions. Although a simheuristic algorithm has been used to obtain garbage collection routes with real-life distances to compare then with the routes of Sabadell, other standard algorithms of the literature could also be tested and compared with them. This work would allow to see the relative difference between several algorithms in a real situation. Moreover, different procedures to generate the initial solution can be tested and their effect on the global performance of the algorithm can be assessed.

In addition, our simheuristic procedure could be extended to consider additional input variables that are typically shaped by some kind of stochastic behaviour, e.g.: waste to be collected or even waste containers themselves. Similarly, the problem setting could be enriched by including historical data to construct a more realistic travel speed model for the study area. An interesting concept in this context is the emerging technique of learnheuristics (Calvet et al., 2017), which complements the simheuristic solving framework by including machine learning techniques to consider problem dynamic inputs –e.g., varying traffic conditions at different times.

Acknowledgements

This work has been partially supported by the Spanish Ministry of Science (Spanish Ministry of Science (PID2019-111100RB-C21 / AEI / 10.13039 / 501100011033, RED2018-102642-T), AGAUR and FEDER (2018 LLAV 00017), and the Erasmus+ programme (2019-I-ES01-KA103-062602). Furthermore, we would like to thank the company SMATSA for providing us with their data and support.

References


A simheuristic algorithm for time-dependent waste collection management...


