

Transformations of node-balanced routing problems

Antonio Martinez-Sykora Tolga Bektaş¹

*Southampton Business School
Centre for Operational Research, Management Science and Information Systems (CORMSIS)
University of Southampton, SO17 1BJ Southampton, United Kingdom*

Abstract

This paper describes a polynomial transformation for a class of unit-demand vehicle routing problems, named node-balanced routing problems (BRP), where the number of nodes on each route is restricted to be in an interval such that the workload across the routes is balanced. The transformation is general in that it can be applied to single or multiple depot, homogeneous or heterogeneous fleet BRPs, and any combination thereof. At the heart of the procedure lies transforming the BRP into a generalized traveling salesman problem (GTSP), which can then be transformed into a traveling salesman problem (TSP). The transformed graph exhibits special properties which can be exploited to significantly reduce the number of arcs, and used to construct a formulation for the resulting TSP that amounts to no more than that of a constrained assignment problem. Computational results on a number of instances are presented.

Keywords. traveling salesman problem, balancing constraints, vehicle routing problems.

1 Introduction

Much ink has been spilt on describing transformations of routing problems into the traveling salesman problem (TSP) in the earlier days, starting with the work of Bellmore and Hong (1977); Hong and Padberg (1977); Jonker and Volgenant (1988) and Rao (1980) for the multiple TSP (mTSP), but only intermittently continued by others for extensions of such problems (see, e.g., GuoXing, 1995; Behzad and Modarres, 2002; Benavent and Martinez-Sykora, 2007). The principle behind these transformations is to exploit the state-of-the-art solution techniques available for the resulting problem after transformation (e.g., the TSP), in order to solve the problem to be transformed (e.g., the mTSP). For further details on transformations of the mTSP, see Bektaş (2006). With the development of problem-specific solution algorithms, however, these transformations have not seen a large take-up. This paper aims to resurrect this line of research by describing a non-intuitive transformation for a problem that has practical significance.

What we name here as the *node-balanced routing problem* (BRP) is a vehicle routing problem (VRP) with unit demands, where the number of nodes visited on each route is bounded by an interval. This is in contrast to the classical setting of the VRP wherein a lower bound, in addition to the generally imposed upper bound (often prescribed to indicate capacity limits on the vehicles), appears as a constraint. If the workload on a route is defined by the number of customers visited on that route, then balancing the workload across the routes can be achieved by tightening the lower and the upper bound. Here, we differentiate between the BRP with single or multiple depots, and with homogeneous or heterogeneous fleet, where the latter differentiation is made with respect to the lower and upper bounds individually imposed on each vehicle. If these bounds are uniform across the fleet, then it is said to be homogeneous. Otherwise the fleet is heterogeneous.

¹Corresponding author.

This problem finds applications in settings where the times spent at customer nodes is significantly more than that of travel, for example home installations or repairs. One real-world application of the BRP is described by Okonjo-Adigwe (1988), which arises in transporting 38 passengers, three of which are wheel-chair users, by using four vehicles one of which is an ambulance and the only vehicle that can carry the wheel-chair users. The vehicles have different capacities, giving way to a heterogeneous BRP. The aim of the problem is to improve the distribution of workload amongst drivers, for which reason a lower bound on the number of capacity usage of each vehicle has been introduced, so that tours carrying very few passengers are avoided in a solution to the problem. A particular and interesting application arises in meter-reading for utility companies, namely the *balanced billing cycle vehicle routing problem* studied by Groër et al. (2009), where, in addition to the primary goal of reducing the total length of the routes, a secondary goal of balancing the routes both in terms of length and the number of visits was considered. The authors solve the problem using a combination of heuristics and integer programming. A more recent and relevant work is by Kritikos and Ioannou (2010), who present a *balanced cargo vehicle routing problem with time windows*, the objective of which is to minimize the number of vehicles, the total distance traveled and to balance the load carried by the vehicles. The latter objective has been modeled by minimizing the sum of the absolute value of the load of each vehicle from an estimated average load. An industrial application of the problem has been presented, which includes 1943 customers and a fleet of 72 vehicles, and which has been solved using a heuristic algorithm. This problem is a generalized version of the BRP in that the loads are general (as opposed to unit) and there are time window restrictions in place.

The BRP is also relevant in routing problems where vehicle or driver utilization is of concern. Imposing a lower bound can avoid solutions where some routes are heavily under-utilized with respect to their capacity. For example, Calvete et al. (2007) study a multi-objective vehicle routing problem with soft time windows where one of the objectives is to minimize the difference between a vehicle's capacity and the total load it carries. We find it pertinent here to quote the following statement, which recognizes the importance of balanced problems and the lack of studies devoted to it:

“Any company that employs multiple drivers to read meters or deliver goods would like each driver to individually work near capacity on each day, implying some degree of balance across the work days. Despite the simplicity and necessity of this objective in a wide variety of scenarios, it appears that only a handful of papers have considered the issue of balance in any detail (Groër et al., 2009).”

Indeed, while traveling salesman and vehicle routing problems have been widely studied (for an excellent introduction, see Laporte, 1992a,b; Toth and Vigo, 2002; Laporte, 2007; Cordeau et al., 2007), the BRP has been much less so. Gouveia et al. (2013) studied the single-depot homogeneous fleet BRP by presenting an integer programming formulation that uses so-called reverse, rounded reverse and enhanced reverse multistar inequalities, which are then used within a Branch-and-Cut algorithm devised to solve the problem. The authors presented computational results for 46 BRP instances with changing bounds. Later, Gouveia and Salazar-González (2013) presented a compact formulation that has the same LP relaxation value as the formulation using enhanced reverse multistar inequalities and a separation algorithm for the same inequalities that runs in polynomial time. Bektaş (2012) described alternative formulations and Benders Decomposition algorithms for the multi-depot homogeneous fleet BRP. To the best of our knowledge, the heterogeneous fleet extension of the problem, either with single or multiple depots, has not yet been investigated, although real-life applications, such as those discussed above, involve such features.

This paper takes a general approach to solve all four variants of the BRP by describing a “one-size-fits-all” transformation of the problem into a Generalized Traveling Salesman Problem (GTSP). The resulting GTSP can then be transformed into a TSP, which in itself poses challenges to solve in terms of the size of the resulting instance. By identifying and exploiting special properties of the transformation graph, we show that the resulting TSP can be formulated as a constrained assignment problem, where the number of additional constraints are no more than the number of clusters in the corresponding GTSP. We also present computational results in solving a number of different types of BRP instances to show the efficiency of transformation.

The rest of the paper is organized as follows. Section 2 first introduces the notation, then describes the transformation for the variants of the problem. We present some special properties of the transformed graph in Section 3. Computational results are presented in Section 4 and conclusions are given in Section 5.

2 Problem description and transformations

A formal description of the BRP is as follows. Let $V = \{1, \dots, n\}$ be the set of nodes (customers) to be visited, each with a demand for a single unit of a commodity. Let $D = \{d_1, d_2, \dots, d_w\}$ be the set of depots where w is the total number of depots. There exist m_k vehicles based at depot d_k where $k \in \{1, \dots, w\}$. $G(V', A)$ is a graph with $V' = V \cup D$ as the set of nodes and $A = \{(i, j) | i \in V', j \in V' \setminus \{(i, j) | d_i \in D, d_j \in D\}\}$ as the set of arcs. For each $(i, j) \in A$, c_{ij} denotes the cost of traveling from node i to node j . We allow either a symmetric or an asymmetric structure on the costs. Let $Q_{min}^{k,r}$ and $Q_{max}^{k,r}$ be the minimum and maximum number of customers to be served by vehicle $r \in \{1, \dots, m_k\}$ of depot $d_k \in D$, respectively. The BRP consists of finding m routes all starting from and ending at their origin depots such that the number of customers visited on each route is between the predefined lower and upper bounds, and each customer is visited exactly once by any vehicle. The objective is to minimize the total cost of travel.

The procedure we propose and describe below transforms the BRP into the GTSP, which is a problem defined on a graph where the set of its nodes is partitioned into clusters, and the problem seeks a minimum cost Hamiltonian path which visits exactly one node from each cluster (Noon and Bean, 1993). In the ensuing exposition, we start by illustrating the transformation for the single depot problem with homogeneous fleet in Section 2.1. Sections 2.2 and 2.3 then show how the transformation applies to the cases with multiple depots and a heterogeneous fleet of vehicles, respectively.

2.1 Single depot and homogenous fleet

This section concerns the BRP with a single depot, i.e., $w = 1$. To simplify the notation, we omit the indices and denote the depot by d and the total number of routes by m . As the fleet is homogeneous, we use Q_{min} and Q_{max} to denote the minimum and maximum number of customers allowed on each route, respectively. We use the notation $\text{BRP}(G, m, Q_{min}, Q_{max})$ to refer the corresponding balanced routing problem defined on graph $G(V', A)$.

We first note that the BRP(G, m, Q_{min}, Q_{max}) has a feasible solution if the following condition holds:

$$\left\lceil \frac{n}{Q_{max}} \right\rceil \leq m \leq \left\lfloor \frac{n}{Q_{min}} \right\rfloor. \quad (2.1)$$

In (2.1), $\left\lceil \frac{n}{Q_{max}} \right\rceil$ ($\left\lfloor \frac{n}{Q_{min}} \right\rfloor$) provides the minimum (maximum) number of vehicles needed to visit all the customers. We assume that the lower and upper bounds on the number of customer per route are tight, i.e., $Q_{min} \geq n - (m - 1)Q_{max}$ and $Q_{max} \leq n - (m - 1)Q_{min}$. If either of these inequalities is not satisfied, say $Q_{min} < n - (m - 1)Q_{max}$, then, this implies $m - 1$ routes are used at maximum capacity and there remain at least Q_{min} customers to be served on remaining route. In this case, Q_{min} can be tightened as $n - (m - 1)Q_{max}$.

The transformation of the BRP(G, m, Q_{min}, Q_{max}) into a GTSP that we propose here results in a graph $G_G(V_G, A_G)$, which has the following set of nodes and arcs:

1. *Clusters and Nodes:* The set of nodes V_G is formed by three sets of clusters, as explained below:

- $2m$ copies of the depot d , denoted by $\{d^1, \dots, d^{2m}\}$, each node in itself forming a (singleton) cluster,
- A dummy ‘‘sink’’ node t connected to the first (d^1) and the last (d^{2m}) copies of the depot, which is in itself a (singleton) cluster and serves as the main depot of the GTSP, and
- n clusters, each holding mQ_{max} copies of node $j \in \{1, \dots, n\}$, and is denoted by the set below:

$$V_G^j = \{j_1^1, \dots, j_1^m, j_2^1, \dots, j_2^m, \dots, j_{Q_{max}}^1, \dots, j_{Q_{max}}^m\}.$$

In this representation, the superscript denotes the number of the route in the BRP solution that node j belongs to, and the subscript denotes its position on the route. In other words, j_q^r is a copy of node $j \in V$ that appears in the q^{th} position on route $r = 1, \dots, m$. Each cluster V_G^j in the transformed graph is constructed in such a way so as to be able to identify the route and the position on the route on which node j is visited in the corresponding BRP solution.

The above construction results in a total of $|V_G| = 2m + 1 + Q_{max}mn$ nodes and in $2m + n + 1$ clusters in the transformed graph. Figure 1 depicts the structure of the resulting graph.

2. *Arcs and Arc Costs:* The transformation results in a graph that is defined with respect to a subset of the whole set of arcs, where the arcs and arc costs are defined as follows. For every $r = 1, \dots, m$, there is an arc from node d^{2r-1} to node i_1^r for each $i \in V$ in the original graph, and there is an arc from each node $i_{Q_{min}}^r, \dots, i_{Q_{max}}^r$ to node d^{2r} . From each node d^{2r} originates an arc, leading to node d^{2r+1} if $r = 1, \dots, m - 1$ or to node t if $r = m$. Finally, the transformed graph includes arcs (i_s^r, j_{s+1}^r) for all $s = 1, \dots, Q_{min} - 1$ and $r = 1, \dots, m$.

In line with the above definitions, all arc costs, except for those that are defined below, are set to ∞ . In the below, h_a denotes the cost of an arc $a \in A_G$. For every $r \in \{1, \dots, m\}$, $q \in \{1, \dots, Q_{max}\}$, we also define $S_q^r = \{1_q^r, \dots, n_q^r\}$ to represent subsets of node copies that are allowed to be visited in the q^{th} position of route r .

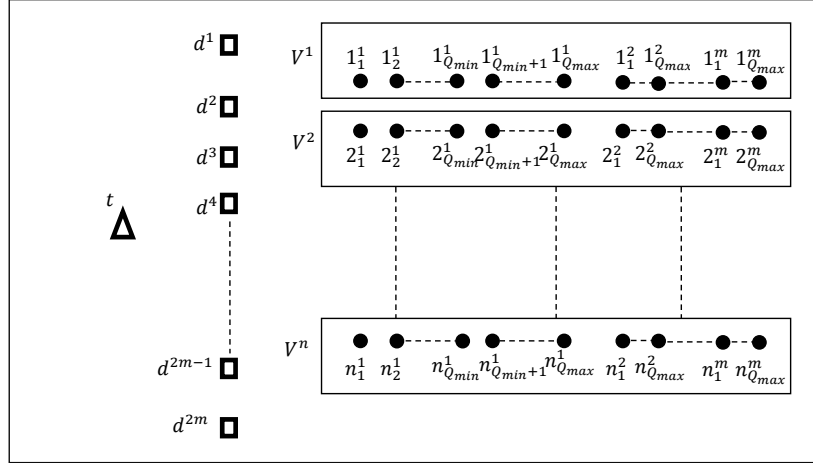


Figure 1: Nodes and clusters of the GTSP after the transformation of a BRP(G, m, Q_{min}, Q_{max}).

- Costs between copies of the depot and other nodes are defined as follows. Arcs (t, d^1) and (d^{2m}, t) have a cost equal to 0. The cost of an arc connecting a copy of a depot and a copy of a customer node is defined as follows,

$$h_{d^{2r-1}, j} = \begin{cases} c_{dj'} & \text{if } j \in S_1^r \\ \infty & \text{otherwise} \end{cases}$$

$$h_{j, d^{2r}} = \begin{cases} c_{j'd} & \text{if } j \in \bigcup_{q=Q_{min}}^{Q_{max}} S_q^r \\ \infty & \text{otherwise,} \end{cases}$$

where $r \in \{1, \dots, m\}$ denotes the index of a route, j' is a customer node of the BRP and j is one of its copies in the GTSP graph. Finally, in order to be able to connect the routes, the cost of arc (d^{2r}, d^{2r+1}) is set equal to 0 for all $r = 1, \dots, m - 1$.

- Intra-customer costs are defined as follows:

$$h_{j,i} = \begin{cases} c_{j'i'} & \text{if } j \in S_q^r, r \in S_{q+1}^r, j' \neq i', r \in \{1, \dots, m\} \text{ and } q \in \{1, \dots, Q_{max} - 1\} \\ \infty & \text{otherwise,} \end{cases}$$

where j' (i') is the customer in the BRP and j (i) one of its copies in the GTSP.

The resulting GTSP after the transformation is denoted by $\text{GTSP}(G_G)$. In what follows we define a correspondence between the BRP(G, m, Q_{min}, Q_{max}) and $\text{GTSP}(G_G)$.

Lemma 2.1 *A feasible solution s_B of the BRP(G, m, Q_{min}, Q_{max}) corresponds to $m!$ solutions of the $\text{GTSP}(G_G)$ with the same cost.*

Proof Let $R_r = \{d, v_1^r, \dots, v_{z_r}^r, d\}$ be a given route of s_B , defined for each $r = \{1, \dots, m\}$, where each route starts and ends at the depot and the number of customers $z_r = |R_r| - 2$ visited on this route is such that $Q_{min} \leq z_r \leq Q_{max}$ and $\sum_{i=1}^m z_r = n$. We will show that route R_r can be represented in the graph G_G in m different ways. This is done by observing that the routes in the GTSP start from node $d^j, j \in \{1, \dots, 2m\}$ with j odd. There are therefore m possibilities to represent route R_r using a path in the GTSP. In the

transformation graph, we assign the number $\frac{j+1}{2}$ to a route that originates at node d^j and ends at d^{j+1} for some $j \in \{1, \dots, 2m\}$, with j odd. For example, route R_1 could originate at node d^1 and end at node d^2 , similarly for route R_2 starting at node d^3 and ending at node d^4 , and more generally route R_r starting at node d^{2r-1} and ending at node d^{2r} . Then, route $R_{\frac{j+1}{2}}$ in s_B can be mapped to a path $P_{\frac{j+1}{2}}^{GTSP}$ in the GTSP as follows:

$$P_{\frac{j+1}{2}}^{GTSP} = \{d^j, v_{1,1}^{r, \frac{j+1}{2}}, \dots, v_{z_r, z_r}^{r, \frac{j+1}{2}}, d^{j+1}\}, \quad (2.2)$$

where the path $\frac{j+1}{2}$ specifies the succession of the nodes visited on this route, on which $v_{1,1}^{i, \frac{j+1}{2}}$ is the first copy of customer v_1^i and $v_{z_r, z_r}^{i, \frac{j+1}{2}}$ denotes the z_r^{th} copy of customer $v_{z_r}^i$ and so on. Then, by construction, route R_r will have the same cost of the cost of path $P_{\frac{j+1}{2}}^{GTSP}$ in the GTSP.

We use this procedure to map each one of the m routes of s_B , which results in m paths in the GTSP in such a way that the total cost of these paths has the same cost as s_B . In order to complete the main route of the GTSP, the m paths are connected using the following arcs:

- Connections between routes use arcs (d^j, d^{j+1}) for all $j \in 2, \dots, 2m - 2$ with j even,
- Connections with the sink cluster t use arcs (t, d^1) and (d^{2m}, t) .

Note that the two sets of arcs above always appear in an optimal solution of the transformed GTSP. Since the m routes used in s_B of the BRP are disjoint, i.e., no customer appears on two different routes, a solution is obtained for the GTSP with the same cost as s_B , where each route of s_B corresponds to path P_r^{GTSP} for some $r = 1, \dots, m$. The fact that r can take m different values corresponding to the odd numbers between 1 and $2m - 1$, yields the result that there exist $m!$ solutions of the GTSP corresponding to s_B . ■

Let us denote by z_{BRP} the cost of the optimal solution of a BRP instance. The proof of Lemma 2.1 implies the following result.

Corollary 2.2 $z_{BRP} \geq z_{GTSP}$, where z_{GTSP} is the cost of the optimal solution of the transformed GTSP solution.

We now present a new lemma establishing the relationship between the cost of the BRP instance and the cost of the GTSP.

Lemma 2.3 $z_{BRP} = z_{GTSP}$.

Proof Let us consider an optimal solution s_G of the transformed GTSP. We first observe that this solution uses arcs (t, d^1) , (d^{2m}, t) and arcs (d^r, d^{r+1}) for $r = 2, \dots, 2m - 2$ with r even, since these are the only arcs visiting such nodes with cost less than ∞ . If these arcs are removed from the solution s_G , the remaining arcs form m paths, each one starting from a node d^r and ending at node d^{r+1} , for $r \in \{1, \dots, 2m - 1\}$ with r' odd. We now build a route R_r for each $r \in \{1, \dots, 2m - 1\}$, r odd, as follows:

1. From copy d^r of the depot, the route in the GTSP must visit copy i_1^r of any customer $i' \in \{1, \dots, n\}$. Therefore, i' is the first customer visited by route R_r .
2. From each node i_q^r , $q = 1, \dots, Q_{min} - 1$, the route must visit another copy i_{q+1}^r of any customer $l \in \{1, \dots, n\}$ for which corresponding customer node $l \in V$ has not previously been visited. Otherwise,

the cluster would be visited more than once and is not a feasible solution of the GTSP. The rest of the arcs that originate from copies i_q^r , $q = 1, \dots, Q_{min} - 1$ have an ∞ cost and therefore will not appear in an optimal solution. At this point, therefore, the first Q_{min} customers of route R_r will have been identified. Let us now denote customer in position Q_{min} visited in R_r by $i'' \in V$.

3. From each customer i_q^r , $q = Q_{min}, \dots, Q_{max} - 1$ the route can visit either another copy i_{q+1}^r of a customer $l \in \{1, \dots, n\}$ as before or the copy of the depot d^{r+1} where the GTSP path corresponding to route R_r would end at.
4. From each customer $i_{Q_{max}}^r$, the route must visit depot d^{r+1} , as any other arc originating from node $i_{Q_{max}}^r$ have an ∞ cost and would therefore not appear in s_G . Note that with this cost structure we guarantee that any path that is formed between copies of depots d^r and d^{r+1} visits at most Q_{max} copies of different customers. Let us denote by $i''' \in V$ the last customer visited in R_j . Then, solution s_G uses the arc (i_q''', d^{r+1}) where $q \in \{Q_{min}, \dots, Q_{max}\}$ so that route R_r is mapped onto path $\{d, i', \dots, i'', \dots, i''', d\}$.

In conclusion, any solution of the GTSP with a cost lower than ∞ can be transformed into a solution of the BRP using the procedure defined above, which shows $z_{BRP} \leq z_{GTSP}$. This result, combined with that of Lemma 2.1, proves $z_{BRP} = z_{GTSP}$. ■

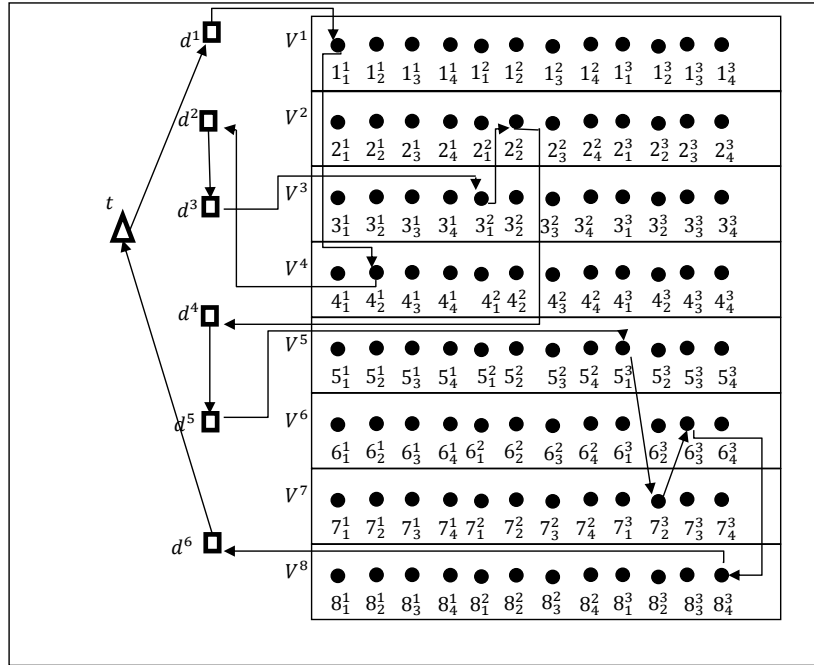
An illustrative example of the transformation is given in Figure 2, where the lower half shows a feasible solution of a BRP instance defined on a graph with eight customers denoted by the node set $\{1, \dots, 8\}$ and a single depot d , on which three routes $R_1 = \{d, 1, 4, d\}$, $R_2 = \{d, 3, 2, d\}$ and $R_3 = \{d, 5, 7, 6, 8, d\}$ are shown. Using the procedure described above, six different solutions for the corresponding GTSP can be built, one of which is shown in the upper half of Figure 2. In this instance, arcs (t, d^1) , (d^2, d^3) , (d^4, d^5) and (d^6, t) have cost equal to 0.

2.2 Multiple depots

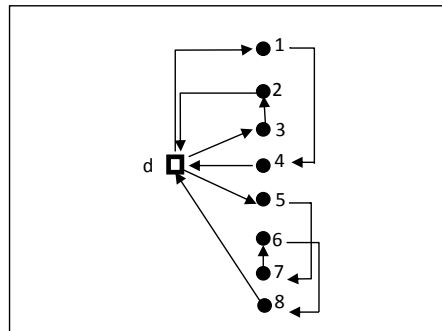
The multiple depot variant of the BRP is denoted as $BRP(G, m_1, \dots, m_w, Q_{min}, Q_{max})$. Similar to the single depot variant, an optimal solution seeks m_i routes originating and returning to each depot $d_i \in D = \{d_1, \dots, d_w\}$ where the number of customers on each route is between Q_{min} and Q_{max} and each customer is visited exactly once.

The transformation is fundamentally the same as in the single depot case and the transformed graph G_G has the same structure as defined above. In what follows we only point out the differences. We also omit any illustrations here since the main idea remains the same. The number of copies of each customer depends on the total number of routes $m_T = \sum_{i=1}^w m_i$ and Q_{max} . Each cluster V^k representing customer k consists of $m_T Q_{max}$ copies. As before, two copies of a depot are needed for each route of that depot, resulting in a total of $2m_T$ copies of depots denoted by $d_1^1, \dots, d_1^{2m_1}, d_2^1, \dots, d_2^{2m_2}, \dots, d_w^1, \dots, d_w^{2m_w}$. Each copy of each depot represents one cluster. The sink node t is again needed in order to construct a single route for the GTSP.

The intra-cost arcs among the customers are the same as in the single depot case. In its general form, let



GTSP



BRP

Figure 2: Example of a transformed solution.

$r \in \{1, \dots, m_i\}$ be the index of a route based on depot d_i . Then:

$$h_{d_i^{2r-1}, j} = \begin{cases} c_{d_i j'} & \text{if } j \in S_1^r \\ \infty & \text{otherwise,} \end{cases}$$

$$h_{j, d_i^{2r}} = \begin{cases} c_{j' d_i} & \text{if } j \in \bigcup_{q=Q_{min}^r}^{Q_{max}^r} S_q^r \\ \infty & \text{otherwise.} \end{cases}$$

In order to connect the routes we consider a cost of 0 for arcs (d_i^{2r}, d_i^{2r+1}) , $r = 1, \dots, m - 1$ and $(d_i^{2m_i}, d_{i+1}^1)$, $i = 1, \dots, w - 1$. The last copy of the last depot $d_w^{2m_w}$ is connected to the sink node t .

Similar results to what is presented in Lemmas 2.1 and 2.3 also apply in this case in that one can show that the optimal value of the BRP($G, m_1, \dots, m_w, Q_{min}, Q_{max}$) is the same as the optimal value of the corresponding GTSP on the transformed graph.

2.3 Heterogenous fleet

This section shows the application of the transformation to the single depot BRP with a heterogenous fleet. In the context of node-balancing, we define the heterogeneity of the fleet relevant to the lower and upper bounds on the number of customers each vehicle is required to visit. In particular, these bounds are prescribed as Q_{min}^r and Q_{max}^r , respectively, for each $r \in \{1, \dots, m\}$. This variant of the problem is denoted by BRP($G, m, Q_{min}^1, \dots, Q_{min}^m, Q_{max}^1, \dots, Q_{max}^m$), where the previously stated objective and the constraints remain the same.

As before, the main structure of the transformation does not change here. In particular, the number of copies made for each customer is $\sum_{r=1}^m Q_{max}^r$, which corresponds to the number of nodes in the cluster representing the copies of that customer node. Then, the first Q_{max}^1 copies of one customer represent the different Q_{max}^1 potential positions of the given customer in the first route R_1 (i.e, the route associate to the first vehicle). Similarly, copies $Q_{max}^1 + 1$ to Q_{max}^2 correspond to those visited by the second vehicle (route R_2), and so on. The total number of copies of depot d is $2m$, i.e, two copies per route. The sink node t also appears here and it has been defined previously. What is different in this case is the set of customers, which requires redefining the cost of the arcs between customers, depots and the intra-customers costs. The structure of the clusters in this case are modified as $V_G^i = \{i_1^1, \dots, i_{Q_{max}^1}^1, i_1^2, \dots, i_{Q_{max}^2}^2, \dots, i_1^m, \dots, i_{Q_{max}^m}^m\}$ for each $i = 1, \dots, n$. The definition of $S_q^r = \{1_q^r, \dots, n_q^r\}$, $r \in \{1, \dots, m\}$, $q \in \{1, \dots, Q_{max}^r\}$ remains as before.

The following costs are defined for each route $r \in \{1, \dots, m\}$,

$$h_{d^{2r-1}, j} = \begin{cases} c_{d j'} & \text{if } j \in S_1^r \\ \infty & \text{otherwise} \end{cases}$$

$$h_{j, d^{2r}} = \begin{cases} c_{j' d} & \text{if } j \in \bigcup_{q=Q_{min}^r}^{Q_{max}^r} S_q^r \\ \infty & \text{otherwise.} \end{cases}$$

The intra-cost arcs among the customers are the same as in the single depot and homogeneous fleet case. After the above modifications, the resulting problem resembles the homogeneous version and the transfor-

mation described in Section 2.1 can be applied. As before, the correspondence between the optimal value of the BRP($G, m, Q_{min}^1, \dots, Q_{min}^m, Q_{max}^1, \dots, Q_{max}^m$) and the optimal value of the corresponding GTSP can be shown in a way that is similar to Lemmas 2.1 and 2.3.

Using a combinations of the procedures defined in Sections 2.2 and 2.3, one can also address the fourth variant of the BRP with multiple depots and heterogenous fleet, the details of which we do not present here. In the next section we present some properties of the transformation graph and a further transformation to the TSP.

3 Properties of the transformation graph

The transformation presented above results in a GTSP, which can then be transformed into the asymmetric TSP. Consequently, any algorithm available for the TSP, exact or heuristic, can be used to solve the transformed problem. However, the number of nodes in the transformed graph might render such an approach impractical. In this section, we will make observations on the special structure that the transformation graph exhibits and show how this particular structure can be exhibited to formulate the resulting TSP with far less number of variables than what would normally be required, and only using a polynomial set of constraints.

Whilst there exist a number of procedures proposed to transform the GTSP into the TSP (see, e.g., Noon and Bean, 1993; Dimitrijević and Šarić, 1997; Laporte and Semet, 1999), we choose the one proposed by Behzad and Modarres (2002) for the very reason that it is, to the best of our knowledge, the only one which does not result in an increase in the number of nodes of the GTSP. As we will show in the below, the original transformation graph G_G already has a significantly large number of nodes, but one which is highly sparse and with a special structure. We now provide more details about the Behzad and Modarres (2002) transformation for the sake of completeness. The principle behind the transformation is to build intra-cluster cycles for each cluster of the GTSP containing two or more nodes with a 0 cost. Let $G_{ATSP} = (V_{ATSP}, A_{ATSP})$ be the graph obtained after applying the Behzad and Modarres (2002) transformation. The set of nodes is the same as in the GTSP, i.e, $V_{ATSP} = V_G$. However, the set of arcs and the costs are modified as follows:

- Clusters containing two or more nodes in G_G are $V_G^i, i = 1, \dots, n$. Successors of nodes i_q^r , for $q = 1, \dots, Q_{max}$ and $r = 1, \dots, m$ for the single-depot homogeneous BRP, for $q = 1, \dots, Q_{max}^r$ and $r = 1, \dots, m$ for the single-depot heterogeneous BRP, and for $r = 1, \dots, m_T$ for the multiple depot BRPs, are defined as follows:
 - If $q < Q_{max}$ (or if $q < Q_{max}^r$ in the case of heterogeneous fleet), then the successor is i_{q+1}^r ,
 - If $q = Q_{max}$ (or if $q = Q_{max}^r$ in the case of heterogeneous fleet) and $r < m$ (or if $r < m_T$ in the case of multiple depots), then the successor is i_1^{r+1} ,
 - If $q = Q_{max}$ (or if $q = Q_{max}^r$ in the case of heterogeneous fleet) and $r = m$ (or if $k = m_T$ in the case of multiple depots), then the successor is i_1^1 .
- The set of arcs among clusters is the same as before. However, the costs are different. If h_{ij} is the cost

of arc $(i, j) \in A_G$, then the cost $c'(i, j)$ of arcs of an arc $(i, j) \in A_{ATSP}$ is defined as follows:

$$c'(i, j) = \begin{cases} 0 & \text{if nodes } i \text{ and } j \text{ belong to the same cluster and } j \text{ succeeds } i, \\ h_{jk} + M' & \text{if } i \text{ and } j \text{ belong to different clusters and } k \text{ succeeds } j, \end{cases} \quad (3.1)$$

where $M' > \sum_{(i,j) \in A_{atsp}} c(i, j)$ is a sufficiently large constant used to make sure that exactly $b = 2m + 1 + n$ arcs in the single depot and $b = 2m_T + 1 + n$ arcs in the multiple depot problem are used in the optimal solution of the transformed ATSP, which correspond to the number of clusters in the GTSP. The optimal solution values are then related as $z_{GTSP} = z_{ATSP} - bM'$, where z_{GTSP} and z_{ATSP} represents the objective function value of an optimal solution of the GTSP and the ATSP, respectively.

Finally, the costs of arcs in the ATSP between singleton clusters and copies of the depot (or depots in case of the multiple depot BRP) and sink node t will be the same as in the GTSP.

The following exposition will, for the sake of simplicity, work on the GTSP obtained from the single depot and homogeneous fleet BRP denoted $BRP(G, m, Q_{min}, Q_{max})$ and use the notation defined in Section 2.1. The transformation results in an ATSP with $2m + 1 + Q_{max}mn$ nodes. The graph of the ATSP transformed from the illustrative instance shown in Figure 2 is presented in Figure 3.

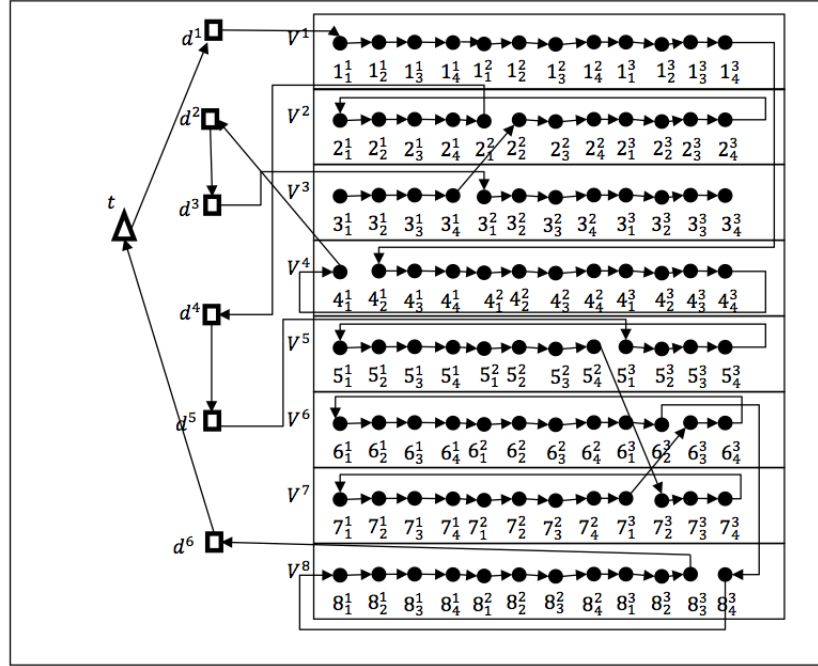


Figure 3: The graph of the final ATSP for the example shown in Figure 2

3.1 An integer programming formulation

Since many arcs in the GTSP graph have an infinity cost, the resulting graph for the ATSP will be highly sparse when such arcs are excluded from further consideration. In an integer programming (IP) formulation of the ATSP, this has implications in significantly reducing the number of binary variables required to

formulate the problem. This is one of our main motivations in proposing to use an integer programming formulation to solve the resulting ATSP, using a binary variable x_{ij} , defined for each $(i, j) \in A_{ATSP}$, which takes the value 1 if arc (i, j) appears in the solution. The formulation is described below and is named IP1:

$$\text{Minimize } \sum_{(i,j) \in A_{ATSP}} c'(i, j)x_{ij} \quad (3.2)$$

subject to

$$\sum_{j \in V_{ATSP} \setminus \{i\}} x_{ij} = 1 \quad \forall i \in V_{ATSP} \quad (3.3)$$

$$\sum_{i \in V_{ATSP} \setminus \{j\}} x_{ij} = 1 \quad \forall j \in V_{ATSP} \quad (3.4)$$

$$\sum_{j, j' \in V_G^i} x_{jj'} = mQ_{max} - 1 \quad \forall i = 1, \dots, n \quad (3.5)$$

$$x_{ij} \in \{0, 1\} \quad \forall (i, j) \in A_{ATSP}. \quad (3.6)$$

The objective function (3.2) minimizes the total cost. Constraints (3.3) and (3.4) guarantee that each node is visited exactly once. Constraints (3.5) are defined for each cluster in the GTSP and avoid subtours, which are tours that are disconnected from the depot, by stating the number of arcs each cluster should contain. Finally, (3.6) forces the variables to be binary.

Constraints (3.5) are different to other types of subtour elimination constraints found in the literature and require further elaboration. Clusters V_G^i defined by mQ_{max} copies of each customer $i = 1, \dots, n$ result in a special structure on the transformed ATSP graph. In particular, there are only few arcs in each cluster V_G^i with cost equal to 0, the rest having a cost equal to ∞ . If the latter arcs are removed from the formulation, then the only possibility for a subtour to form would be within arcs of cost 0. Equalities (3.5) serve to eliminate such intra-cluster subtours by setting the total number of arcs that should be visited within each cluster in a feasible solution to $mQ_{max} - 1$, i.e., they force the formation of intra-cluster paths.

Note that since there are several arcs that would appear in our GTSP graph, the following binary variables will take the value one in all the solutions of the ATSP with an optimal cost:

- $x_{td^1} = 1$ and $x_{d^{2m}t} = 1$,
- $x_{d^{2r}d^{2r+1}} = 1, r = 1, \dots, m - 1$.

We now show that a solution obtained by the formulation IP1 does not contain any subtours.

Lemma 3.1 *An optimal solution obtained by IP1 does not contain subtours formed between clusters and is therefore an optimal solution of the corresponding ATSP.*

Proof We first observe that, due to the sparsity of the graph G_{ATSP} , no subtour can be formed between the singleton clusters defined by each node t, d^1, \dots, d^{2m} . Therefore, any possible subtour would only be formed within the nodes corresponding to clusters V_G^i or between copies of customers and depots. In the below, we will first prove that any subtour in which node t appears must visit all the copies of the depots,

and then show that no subtour can be formed within clusters $V_G^i, i = 1, \dots, n$. All the nodes of the ATSP will be visited in the same subtour, which will be the TSP tour itself.

Since $x_{td^1} = 1$ and $x_{d^{2m}t} = 1$ in any feasible solution, then d^1 and d^{2m} have to be in the same subtour as node t . By construction, the only nodes reachable from node d^1 are given by the set $R(d^1) = \{i_1^1 \mid i = 1, \dots, n\}$. All nodes reachable from any node $i \in R(d^1)$ are $R(i) = \{j_2^1 \mid j = 1, \dots, n, j \neq i\}$. We denote by $R(i_1^k) = \{j_{k+1}^1 \mid j = 1, \dots, n, j \neq i\}$ where $k = 1, \dots, Q_{min} - 1$. From copy number $i_{Q_{min}}^1$ to copy number $i_{Q_{max}-1}^1$ the set of reachable nodes is $R(i_1^k) = \{i_{k+1}^1 \mid i = 1, \dots, n\} \cup \{d^2\}$. In order to define the first route, the only node reachable from nodes $i_{Q_{max}}^1, i = 1, \dots, n$ is the node d^2 . Therefore, the only arcs that can be used to leave the nodes defining the first route, $i_k^1, i = 1, \dots, n$ and $k = 1, \dots, Q_{max}$, have node d^2 as their destination. By construction of the graph, there do not exist arcs between any node $i_k^1, i = 1, \dots, n, k = 1, \dots, Q_{max}$ and any node $i_k^j, i = 1, \dots, n, k = 1, \dots, Q_{max}$ and $j = 1, \dots, m, j \neq 1$. Therefore, node d^2 must belong to the same subtour as node t . We now observe that as $x_{d^{2r}d^{2r+1}} = 1, r = 1, \dots, m - 1, d^3$ will succeed d^2 in any solution, and d^3 will be the starting node of the second path. By applying the same procedure in an iterative fashion, it can be seen that all copies of the depots d^1, \dots, d^{2m} appear in the same subtour in which the sink node t also appears.

We now assume that there exists a subtour between clusters $V_G^i, i = 1, \dots, n$. Given the structure of the graph, the only possibility for a subtour to appear would be within copies of customers belonging to the same path in the GTSP. Let $k \in \{1, \dots, m\}$. The only way to visit any customer of the set $K^1 = \{i_k^1 \mid i = 1, \dots, n\}$ is by using an arc originating from node d^{2k-1} , so no node from K^1 would be included in the subtour. In addition, nodes from $K^2 = \{i_k^2 \mid i = 1, \dots, n\}$ can be visited by using only arcs whose predecessor is a node from K^1 , so no node of K^2 would be in the subtour. By proceeding in a similar and an iterative fashion, it can be seen that no subtour between clusters $V_G^i, i = 1, \dots, n$ can be built, implying that no subtour will appear in any solution of IP1. ■

The next property we present is used to reduce the number of binary variables in IP1.

Lemma 3.2 *Let IP2 denote the formulation IP1 where constraints (3.6) are replaced by the following:*

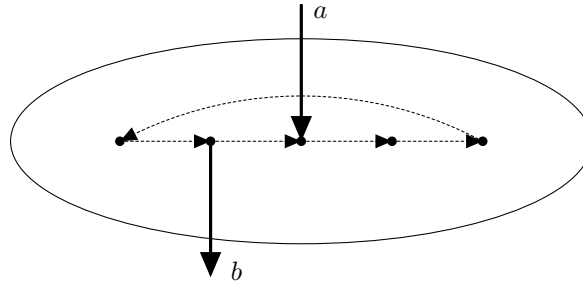
$$x_{i'j'} \in \{0, 1\} \quad \forall (i', j') \in A_{ATSP}, i' \in V_G^i, j' \in V_G^j, i \neq j \quad (3.7)$$

$$0 \leq x_{i'j'} \leq 1 \quad \forall (i', j') \in A_{ATSP}, i' \in V_G^i, j' \in V_G^j, i = j. \quad (3.8)$$

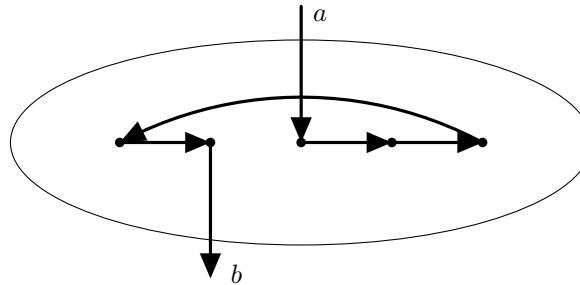
Then, there is one-to-one correspondence between the set of integer solutions of formulations IP1 and IP2.

Proof Proof of Lemma 3.2 Since any cluster $V_G^i, i = 1, \dots, n$ is visited exactly once, then there exist only two arcs incident to any node of the cluster. In the presence of equalities (3.5), there can only be $mQ_{max} - 1$ arcs in the cluster, which enforces the formation of a path within a cluster. All arcs on the path will therefore attain the value of 1 in the solution, regardless of whether there is integrality requirements on the corresponding arc variables. Therefore, there exist no fractional solutions satisfying equalities (3.5). ■

An example to Lemma 3.2 is given in Figure 4, where the dashed arcs in Figure 4(a) show all available arcs within a given cluster and Figure 4(b) shows the only possible path forming within the cluster if arcs a and b are in the solution.



(a) All available arcs within a cluster



(b) The only path forming within the cluster

Figure 4: Arcs in cluster G_G^i

The difference between IP1 and IP2 is that the integrality requirements on some binary variables of IP1 have been relaxed, yielding a formulation that has a smaller number of binary variables. The effect of the reduction of the number of variables will be shown through computational experiments in the next section.

4 Computational experiments

This section presents the results of computational experiments conducted on the four variants of the BRP and in solving the formulations IP1 and IP2 using CPLEX version 12.5.0.0 under a time limit of three hours. We note that the performance of the formulation largely depends on the optimization package and the settings used, the investigation of which is beyond the remit of this paper. Instead, the experiments presented here are aimed at illustrating the generality of the transformation and are indicative of the performance of the resulting formulation. For benchmarking purposes, we compare the results obtained by solving formulations IP1 and IP2 with those obtained by solving a polynomial-size single commodity (based) formulation, denoted by FF, defined for each of the four variants of the BRP. As these formulations are not central to the paper, we present them in the appendix.

The BRP instances tested here have been derived from two symmetric TSPLIB instances and two asymmetric instances available at <http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/>, namely gr24 and eil51, 24-node and 51-node symmetric instances, respectively, and ftv35 and ftv64, 36-node and 65-node asymmetric instances respectively. We present the results separately for each variant of the BRP.

4.1 Single depot and homogenous fleet

Sixteen symmetric and 14 asymmetric single depot and homogeneous fleet BRP instances were generated. The format $\text{name}n-r-Q_{\min}-Q_{\max}$ has been adopted in naming the BRP instances, where name is the name of the generating TSPLIB instance, n is the number of nodes and r denotes the number of routes. The first node of each generating TSP instance represents the depot in the resulting BRP.

In this variant of the problem, when all vehicles are given the same bounds and all the routes originate from a single depot, the transformation graph exhibits a high degree of symmetry. In particular, any two routes in an optimal solution can be swapped, resulting in an alternative optimal solution (see Lemma 2.1). In order to prohibit such symmetries, and to improve the computational efficiency of the model, we have explored two sets of what we call symmetry breaking inequalities, presented below.

We assume that the vehicles are sorted in an arbitrary order with respect to their index. The first set of inequalities, denoted by BS1, forces the first vehicle to visit at least the same number of customers as the second vehicle. A similar inequality can be written for the second and third vehicles. One can continue iteratively in such a fashion, until an inequality with the penultimate and the last vehicle is written. The last vehicle in the order will therefore be restricted to visit the lowest number of customers. These inequalities in the can be written for the transformation graph as follows:

$$\sum_{j,j' \in T_r} x_{jj'} \geq \sum_{j,j' \in T_{r+1}} x_{jj'}, \quad \forall r \in \{1, \dots, m-1\}, \quad (4.1)$$

where $T_r = \left(\bigcup_{i=(r-1)Q_{\max}+1}^{rQ_{\max}-1} S_i^r \right) \cup S_1^{r'}$, $r \in \{1, \dots, m\}$, and r' being the successor route from r . We assume that if $r = m$ then $r' = 1$. The second set of inequalities, denoted by BS2, restrict some of the customers to be visited only by certain vehicles. Initially we force the first customer to be visited by the first vehicle, i.e., $\sum_{j,j' \in W_1^1} x_{jj'} \leq Q_{\max} - 1$ where $W_i^r = \bigcup_{j \in V_G} S_j^r$. Since the second customer can be visited by the first vehicle as well, we restrict the second customer to be visited by either the first or the second vehicle. Continuing in an iterative fashion, we restrict customer t , where $t \leq m-1$, to be visited by the first t vehicles. These inequalities can be written in a general way as follows, which take into account the number of arcs in a subset of nodes (not in the cut).

$$\sum_{i=1}^t \sum_{i'=i}^t \sum_{j,j' \in W_i^{i'}} x_{jj'} \leq tQ_{\max} - 1, \quad \forall t \in \{1, \dots, m-1\}. \quad (4.2)$$

Table 1 presents the results. The second column shows the number of nodes in the resulting ATSP after the transformation. The optimal value of the TSP for each instance is shown in the third column to serve as a lower bound. Columns four and five show the final optimality gap and the solution time of FF. Columns six to ten respectively show, for IP1, the final optimality gap (in percent), the value of an optimal or the best known solution, computation time (in seconds), the number of binary variables in the formulation, and the number of binary variables in IP1 as a percentage of what would be needed to model a full transformed graph. Columns eleven to fifteen show the same information for IP2. The last four columns present similar statistics for IP1 supplemented by BS1 or B2. We did not test IP2 combined with BS1 or BS2 as its performance is generally dominated by IP1. The table also shows average gaps and times, and the number of optimally solved instances (row #opt), separately for symmetric and asymmetric instances.

Table 1: Computational results for the single depot and homogeneous fleet BRP

Name of the instance	Nodes		FF		IP1		IP2		IP1+BS1		IP1+BS2		
	LB	GAP	Time (sec)	GAP	UB	Time (sec)	%	BV	%	BV	%	GAP	Time (sec)
<i>Symmetric instances</i>													
gr24-2-5-18	833	1272	0	1	1442	27.29	18630	2.68	17895	2.58	0	31	27.85
gr24-2-7-16	741	1272	0	0.91	1448	18.17	16330	2.97	15687	2.86	0	21.08	13.95
gr24-2-10-13	603	1272	0	9.52	1489	15.14	12880	3.54	12375	3.4	0	16.51	24.41
gr24-2-11-12	557	1272	0	13.25	1499	11.53	11730	3.78	11271	3.63	0	12.16	7.3
gr24-3-4-10	697	1272	0	4.95	1765	20.97	14766	3.04	14215	2.93	0	24.43	19.5
gr24-3-5-9	628	1272	0	7.95	1811	24.66	13041	3.31	12559	3.18	0	29.04	28.27
gr24-3-6-9	628	1272	0	15.67	1814	21.54	12972	3.29	12490	3.17	0	25.73	27.09
gr24-3-7-8	559	1272	0	10.26	1850	17.67	11247	3.6	10834	3.47	0	21.39	18.18
el151-6-5-11	3313	426	0	6433.8	561	10800	152100	1.39	149401	1.36	1.45	10800	11.4
el151-6-6-10	3013	426	1.08	10800	567	1565.7	136500	1.5	3806	134101	1.48	1865.8	0
el151-6-7-9	2713	426	0	3358.8	577	4270.8	120900	1.64	577	4884.8	118501	1.61	0
el151-6-8-9	2713	426	2.01	10800	583	3164.7	120600	1.64	583	2944.3	118501	1.61	0
el151-8-3-10	4017	426	0	7446.2	621	10800	183200	1.13	2.01	621	10800	180001	1.12
el151-8-4-9	3617	426	-	10800	1.2	636	10800	1.24	1.3	636	10800	159601	1.22
el151-8-5-8	3217	426	2.07	10800	0	651	1966.9	141600	1.37	0	651	1640.9	139201
el151-8-6-7	2817	426	4.61	10800	0.38	681	10800	1.52	0.41	681	10800	118801	1.49
Average			0.65	4456.39	0.27	3395.32	120800	1.52	0.33	3539.69	0.42	3542.34	0.18
#opt			11		12				12		12		14
<i>Asymmetric instances</i>													
fv35-2-5-30	2105	1473	0	0.81	1489	453.39	72870	1.65	1489	375.66	70911	1.6	0
fv35-2-10-25	1755	1473	0	6.01	1491	224.8	59920	1.95	1491	173.81	58311	1.89	0
fv35-2-15-20	1405	1473	0	4.31	1491	70.63	46970	2.38	1491	185.25	45711	2.32	0
fv35-2-17-18	1265	1473	0	27.86	1514	152.23	41790	2.61	1514	154.78	40671	2.54	0
fv35-3-5-25	2362	1473	0	2.11	1541	359.42	90405	1.62	1541	544.04	87991	1.58	0
fv35-3-8-19	2002	1473	0	2.25	1557	206.78	67410	1.68	1583	236.28	65626	1.64	0
fv35-3-10-15	1582	1473	0	4.84	1583	139.05	52080	2.08	1583	118.72	50716	2.03	0
fv35-3-11-13	1372	1473	0	3.51	1586	134.11	47179	2.51	1586	123.52	43261	2.3	0
fv64-7-7-20	8975	1839	5.15	10800	4.42	2393	551040	0.68	4.54	2396	10800	551040	0.68
fv64-7-8-12	5391	1839	2.88	10800	2.32	2411	10800	317632	1.09	1.8	2411	10800	313153
fv64-7-9-10	4495	1839	1.81	10800	2.81	2500	258944	1.28	3.22	2506	10800	255361	1.26
fv64-9-5-11	6355	1839	3.38	10800	0	2588	9890.23	372672	0.92	0	2588	10135	367489
fv64-9-6-10	5779	1839	5.28	10800	3.1	2711	10800	334656	1	2.5	2702	10800	330049
fv64-9-7-8	4627	1839	0	10049	0.95	2772	10800	259200	1.21	0.73	2772	10800	255745
Average			1.32	4578.62	0.97	4687.90		1.21	0.91	4717.65	1.08	4666.24	0.69
#opt			9		9				9		9		11

Table 1 shows that all symmetric instances with 24 nodes are optimally solved by all formulations, where IP1 and IP2 yield a similar performance. As for the symmetric instances with 51 nodes, FF was able to optimally solve two instances with six routes and one with eight routes, whereas IP1, IP2 and BS1 have been able to solve only four instances to optimality, three with six routes and one with eight routes. BS2 in this case produces better results both in terms of the number of optimally solved instances, which is six, and the overall optimality gap. Similar results hold for asymmetric instances, in particular, all instances with 36 nodes are solved to optimality by all formulations, but most of the 65-node instances remain unsolved. In particular, FF, IP1, IP2 and BS1 can only optimally solve one 65-node instance each. In this case, BS2 outperforms all other formulations by solving three 65-node instances to optimality and yielding the least average optimality gap.

In general, the computational results obtained by both IP1 and IP2 are worse than the results obtained by solving FF when the interval $[Q_{min} Q_{max}]$ is wide. In addition, there is no improvement on using IP2 instead IP1. However, the computational efficiency of IP1 and IP2 becomes more apparent as the interval tightens, i.e., when the problem is better balanced. In particular, IP1 or IP2 are more efficient in solving instances with more nodes and more restrictive routes such as eil51-6-8-9, eil51-8-5-8, ftv64-9-5-11, either in providing optimal solutions or a yielding a tighter optimality gap.

Table 1 also shows the impact of using inequalities BS1 and BS2 in IP1. Interestingly, inequalities BS1 do not help in improving the computational times, and for some instances the performance worsens. This can be explained by the density of the inequalities, which increase the complexity of solving the formulations, leading to slower performance. In contrast, inequalities BS2 clearly improve the results, both with respect to computational time and the number of instances solved to optimality.

4.2 Multiple depots and homogenous fleet

Eight BRP instances have been generated with $w = 2$ or $w = 4$ depots and a homogeneous fleet. The first w nodes of the generating TSP instances serve as the depots. Table 2 describes the characteristics of these instances in more detail. The second column of the table shows the number of depots in each instance. Columns m_1 – m_4 represent the number of vehicles at (and therefore the routes that originate from) depots d_1 – d_4 , respectively. The last column shows the total number of routes.

Table 2: Properties of the instances generated for the multiple depot and homogeneous fleet BRP

Name of the instance	w	m_1	m_2	m_3	m_4	Total
gr24md.5.10	2	2	1	-	-	3
gr24md.6.8	2	2	1	-	-	3
gr24md.3.7	2	2	2	-	-	4
gr24md.4.6	2	2	2	-	-	4
eil51md.8.15	4	1	1	1	1	4
eil51md.10.13	4	1	1	1	1	4
eil51md.6.10	4	2	2	1	1	6
eil51md.7.9	4	2	2	1	1	6

Table 3 presents the computational results using the same format as Table 1.

The results presented in Table 3 show that all eight instances but one with 51 nodes are solved to optimality within the allowed time limit by either IP1 or IP2. One interesting observation is that, although gr24md.6.8

Table 3: Computational results for the multiple depot and homogeneous fleet BRP

Instances	Nodes	FF		IP1					IP2				
		GAP	Time(sec)	GAP	UB	Time(sec)	BV	%	GAP	UB	Time(sec)	BV	%
gr24md_5.10	667	0.00	101.35	0.00	1494	73.91	13600	3.06	0.00	1494	50.02	12940	2.91
gr24md_6.8	535	0.00	75.35	0.00	1508	59.55	10498	3.67	0.00	1508	20.20	9970	3.49
gr24md_3.7	625	0.00	313.73	0.00	1620	36.09	12237	3.14	0.00	1620	41.99	11621	2.98
gr24md_4.6	537	0.00	490.14	0.00	1620	5.74	10125	3.52	0.00	1620	8.64	9597	3.33
eil51md_8.15	2829	10.97	10800	2.54	444	10800	125589	1.57	2.54	444	10800	122649	1.53
eil51md_10.13	2453	17.29	10800	0.00	448	4585.68	107165	1.78	0.00	448	5652.87	104617	1.74
eil51md_6.10	2833	14.27	10800	0.00	486	2972.66	121267	1.51	0.00	486	1905.90	118327	1.47
eil51md_7.9	2551	18.79	10800	0.00	500	1825.66	107449	1.65	0.00	500	1586.51	104803	1.61
Average		7.67	5522.57	0.32		2544.91					2508.27		
#opt		4		7					7				

has tighter bounds than gr24md_5.10, it required less time to be solved to optimality, both with IP1 and IP2. A similar behaviour can be observed between instances gr24md_3.7 and gr24md_4.6. These results once again indicate that the transformation works better when the difference between Q_{min} and Q_{max} is small. FF, on the other hand, failed to solve all instances with 51 nodes and has yielded optimality gaps that are more than 10%. The transformation in this case leads to formulations which produce better results than FF.

Further to the results in Table 3, we have also solved some of the instances described in Bektaş (2012) who describes a specialized decomposition algorithm, denoted BD, for this variant of the problem. The comparison results are shown in Table 4 using the same format as before, with the exception of the last two columns which show the optimal solution value (or the best upper bound) reported by Bektaş (2012) and the computational time required by (or percentage optimality gap yielded by the best algorithm) within three hours.

Table 4: Computational results for the multiple depots and homogeneous fleet BRP

Instances	Nodes	FF		IP1					IP2					BD	
		GAP	Time(sec)	GAP	UB	Time (sec)	BV	%	GAP	UB	Time (sec)	BV	%	Opt/UB	Best T/GAP
ftv33	1033	0.00	1427.01	0.00	1579.00	161.05	29061	2.73	0.00	1579.00	130.63	28069	2.63	1579.00	1507.05
ftv35	1369	0.00	1695.50	0.00	1669.00	121.28	42709	2.28	0.00	1669.00	78.53	41389	2.21	1669.00	75.87
ftv38	1489	0.00	319.34	0.00	1730.00	155.62	50177	2.26	0.00	1730.00	162.92	48737	2.20	1730.00	319.34
p43	2469	0.31	10800	62.59	5695.00	10800	96273	1.58	62.91	5702.00	10800	93813	1.54	5695.00	0.31%
ftv44	2589	0.00	738.18	0.00	1802.00	1149.34	105785	1.58	0.00	1802.00	1446.43	103265	1.54	1802.00	97.42
ftv47	3689	0.00	875.98	3.17	1975.00	10800	164133	1.21	3.75	1975.00	10800	160533	1.18	1975.00	167.53
ry48p	2769	0.00	8734.62	0.00	15864.00	1447.42	120893	1.58	0.00	15864.00	407.31	118133	1.54	15864.00	3481.26
ftv53	3069	0.00	1902.67	0.00	7396.00	5865.72	148109	1.57	5.20	7536.00	10800	145049	1.54	7396.00	1038.64
ftv55	3193	1.86	10800	0.00	2013.00	426.31	154237	1.51	0.00	2013.00	195.92	151117	1.48	2013.00	5664.32
Average		0.24	4143.70	7.31		3436.30			7.98		3869.08				
#opt		7		7					6						

As the results in Table 3 show, by using IP1 and IP2, the computational time to optimally solve instances ftv33, ftv38, ry48p and ftv55 is considerably reduced. Only for one instance, namely ftv47, neither IP1 nor IP2 were able to prove optimality within three hours; in this case the best solution obtained by BD is an optimal solution.

4.3 Single depot and heterogeneous fleet

Sixteen single depot BRP instances with a heterogeneous fleet have been generated based on the same four TSPLIB instances mentioned above, with the first node considered as the depot. Table 5 describes the instances in more detail, where the second column n_r shows the number of vehicles (routes), columns

R_1 – R_7 show the Q_{min} and Q_{max} values for routes 1–6, respectively, for each instance. The instances are generated in such a way the second and the fourth instance generated from each instance type are more restrictive than the first and the third instances respectively. Computational results for these instances are presented in Table 6 in the same way as the previous tables.

Table 5: Properties of the instances generated for the single depot and heterogeneous fleet BRP

Name of the instance	nr	R_1	R_2	R_3	R_4	R_5	R_6	R_7
gr24hf.1	2	6–10	13–17	-	-	-	-	-
gr24hf.2	2	7–9	14–16	-	-	-	-	-
gr24hf.3	3	4–7	6–10	8–12	-	-	-	-
gr24hf.4	3	5–6	7–8	9–10	-	-	-	-
eil51hf.1	5	3–5	11–15	8–10	11–13	13–13	-	-
eil51hf.2	5	4–4	12–13	8–9	12–12	13–13	-	-
eil51hf.3	6	4–6	3–5	7–11	10–12	11–13	12–14	-
eil51hf.4	6	5–5	4–5	7–9	10–11	11–11	12–12	-
ftv35hf-1	3	5–9	11–14	14–18	-	-	-	-
ftv35hf-2	3	6–8	12–14	15–17	-	-	-	-
ftv35hf-3	4	3–7	6–9	9–12	12–14	-	-	-
ftv35hf-4	4	4–6	6–8	10–12	12–14	-	-	-
ftv64hf-1	6	3–5	6–9	10–13	12–15	14–17	15–18	-
ftv64hf-2	6	4–5	6–7	10–11	12–13	15–15	15–16	-
ftv64hf-3	7	2–4	4–6	5–7	8–11	12–14	14–16	15–19
ftv64hf-4	7	2–3	4–5	5–5	10–10	12–12	15–15	15–16

Table 6: Computational results for the single depot and heterogeneous fleet BRP

Instances	Nodes	FF		IP1				IP2					
		GAP	Time(sec)	GAP	UB	Time(sec)	BV	%	GAP	UB	Time(sec)	BV	%
<i>Symmetric instances</i>													
gr24hf.1	626	0.00	1.95	0.00	1448	30.00	13550	3.46	0.00	1448	26.73	12929	3.3
gr24hf.2	580	0.00	3.92	0.00	1448	10.15	12400	3.69	0.00	1448	8.38	11825	3.52
gr24hf.3	674	0.00	299.98	0.00	1756	40.99	14218	3.13	0.00	1756	40.67	13551	2.99
gr24hf.4	559	0.00	62.35	0.00	1788	13.09	11389	3.65	0.00	1788	10.87	10837	3.47
eil51hf.1	2811	- [†]	10800	0.00	518	3404.83	128250	1.62	0.00	518	2192.63	125951	1.56
eil51hf.2	2561	28.60	10800	0.00	525	3247.57	115350	1.76	0.00	525	975.42	113301	1.73
eil51hf.3	3013	22.93	10800	0.00	545	1414.12	135950	1.5	0.00	545	569.33	133551	1.47
eil51hf.4	2613	-	10800	0.00	556	1244.44	115450	1.7	0.00	556	567.46	113451	1.66
Average #opt		8.59 [‡]	5446.03	0.00		1175.65			0.00		548.94		
		4		8					8				
<i>Asymmetric instances</i>													
ftv35hf.1	1442	0.00	47.68	0.00	1541	59.14	47040	2.26	0.00	1541	57.37	45816	2.2
ftv35hf.2	1372	0.00	26.76	0.00	1550	95.41	44415	2.36	0.00	1550	43.62	43261	2.3
ftv35hf.3	1479	0.00	224.74	0.00	1631	68.66	47110	2.16	0.00	1631	68.87	45921	2.1
ftv35hf.4	1409	0.00	122.39	0.00	1636	27.79	44520	2.24	0.00	1636	89.83	43401	2.19
ftv64hf.1	4941	3.11	10800	0.00	2156	3481.81	292288	1.2	0.00	2156	4279.93	288129	1.18
ftv64hf.2	4301	19.86	10800	0.00	2180	2230.37	250560	1.35	0.03	2180	10800	247041	1.34
ftv64hf.3	4943	-	10800	0.00	2205	2214.99	288256	1.18	0.00	2205	2281.81	284225	1.16
ftv64hf.4	4239	-	10800	0.00	2223	2334.30	242304	1.35	0.00	2223	2549.40	238977	1.33
Average #opt		3.83 [‡]	5452.70	0.00		1314.06			0.00		2521.35		
		4		8					7				

[†]: Feasible solution not found within the time limit.

[‡]: Average values calculated only over instances that are either solved to optimality or for which a feasible solution was found.

The results presented in Table 6 are consistent with those presented earlier in that the efficiency of the transformation increases as the width of the interval defined by $[Q_{min}, Q_{max}]$ is reduced. All 24-node instances and 36-node instances are solved by any of the three formulations, whereas FF failed to optimally solve any of the 51-node and 65-node instances. In contrast, IP1 and IP2 optimally solved all the 51-node instances, and IP1 additionally solved all the 65-node instances to optimality.

4.4 Multiple depots and heterogeneous fleet

The last set of computational experiments we present here concern the most general variant of the BRP, namely that with multiple depots and heterogeneous fleet. Sixteen representative instances have been generated for this case, for which details are given in Table 7. The second column of this table shows the number of depots for each instance, which in this case is set to two for all instances. The remaining columns define Q_{min} and Q_{max} for each vehicle at each depot. For example, eil51c.4 requires three vehicles to be dispatched from each of the two depots. The first vehicle (route) of the first depot needs to visit between three and six customers, the second vehicle has to visit between seven and ten customers and the last vehicle of the first depot has to visit between nine and twelve customers. The second depot has three vehicles, for which the respective bounds are [6,8], [7,12] and [11,15] on routes R_1 , R_2 and R_3 . Table 8 shows the computational results obtained on these instances.

Table 7: Properties of the instances generated for the multiple depot and heterogeneous fleet BRP

Name of the instance	ndep	d_1					d_2				d_3			
		routes	R_1	R_2	R_3	routes	R_1	R_2	R_3	routes	R_1	R_2	R_3	
gr24c.1	2	1	5-9	-	-	1	13-17	-	-	-	-	-	-	
gr24c.2	2	1	7-7	-	-	1	15-15	-	-	-	-	-	-	
gr24c.3	2	2	4-6	6-8	-	2	4-6	6-8	-	-	-	-	-	
gr24c.4	2	2	4-5	6-7	-	2	4-5	6-7	-	-	-	-	-	
eil51c.1	2	2	6-10	10-14	-	2	8-12	15-20	-	-	-	-	-	
eil51c.2	2	2	6-10	10-14	-	3	4-7	8-12	15-20	-	-	-	-	
eil51c.3	2	3	3-6	7-10	9-12	2	10-13	15-20	-	-	-	-	-	
eil51c.4	2	3	3-6	7-10	9-12	3	6-8	7-12	11-15	-	-	-	-	
ftv35c.1	2	2	5-9	11-14	-	1	14-18	-	-	-	-	-	-	
ftv35c.2	2	2	6-8	11-13	-	1	14-16	-	-	-	-	-	-	
ftv35c.3	2	2	3-7	6-9	-	2	9-12	12-14	-	-	-	-	-	
ftv35c.4	2	2	4-6	6-8	-	2	9-10	12-13	-	-	-	-	-	
ftv64c.1	2	3	3-5	6-9	10-13	3	12-15	14-17	15-18	-	-	-	-	
ftv64c.2	2	3	4-5	6-7	10-11	3	12-13	14-14	15-16	-	-	-	-	
ftv64c.3	3	3	2-3	3-4	4-6	3	5-7	6-10	8-11	2	12-14	13-15	-	
ftv64c.4	3	3	2-2	3-3	4-5	3	6-6	6-8	9-10	2	14-14	15-15	-	

Table 8: Computational results for the multiple depots and heterogeneous fleet BRP

Instances	Nodes	FF		IP1					IP2				
		GAP	Time(sec)	GAP	UB	Time(sec)	BV	%	GAP	UB	Time(sec)	BV	%
<i>Symmetric instances</i>													
gr24c.1	577	0.00	99.63	0.00	1308	44.02	11927	3.59	0.00	1308	28.26	11355	3.42
gr24c.2	489	0.00	113.09	0.00	1314	13.43	9815	4.11	0.00	1314	13.57	9331	3.91
gr24c.3	625	0.00	1941.76	0.00	1620	11.17	12061	3.09	0.00	1620	10.24	11473	2.94
gr24c.4	537	0.00	2002.81	0.00	1620	7.34	10037	3.49	0.00	1620	4.48	9533	3.31
eil51c.1	2753	- [†]	10800	3.71	473	10800	126278	1.67	6.05	483	10800	123702	1.63
eil51c.2	3098	73.95	10800	3.67	495	10800	140979	1.47	4.81	503	10800	138081	1.44
eil51c.3	3000	21.5	10800	0.00	494	6999.07	136030	1.51	0.00	494	7740.27	133285	1.48
eil51c.4	3100	-	10800	0.00	520	2000.95	138726	1.44	0.00	520	3203.59	135891	1.41
Average #opt		15.91 [‡]	5919.66	0.92		3834.50			1.36		4075.05		
		4		6					6				
<i>Asymmetric instances</i>													
ftv35c.1	1401	0.00	1436.39	0.00	1516	29.73	44404	2.26	0.00	1516	36.04	43215	2.2
ftv35c.2	1231	0.00	907.31	0.00	1529	25.83	39610	2.62	0.00	1529	30.58	38557	2.55
ftv35c.3	1437	0.00	2942.61	0.00	1653	70.81	44472	2.16	0.00	1653	77.56	43317	2.1
ftv35c.4	1369	0.00	2543.42	0.00	1658	20.46	38488	2.06	0.00	1658	20.26	37503	2
ftv64c.1	4864	-	10800	0.00	2201	409.26	283248	1.20	0.00	2201	2345.17	280421	1.19
ftv64c.2	4171	-	10800	0.00	2234	2994.79	238833	1.37	0.00	2234	3711.76	235432	1.35
ftv64c.3	3425	-	10800	0.00	2291	6139.37	215946	1.84	0.00	2291	3584.4	198401	1.69
ftv64c.4	3921	-	10800	0.00	2291	6113.86	215946	1.40	0.00	2291	10710	212908	1.39
Average #opt		0.00 [†]	6378.72	0.00		1975.51			0.00		2564.47		
		4		8					8				

[†]: Feasible solution not found within the time limit.

[‡]: Average values calculated only over instances that are either solved to optimality or for which a feasible solution was found.

The results shown above parallel those presented earlier, particularly with respect to the quality of the solutions and the computational time. These results also indicate that, despite the increased difficulty of the problem with the introduction of multiple depots and heterogeneous fleet, the transformed formulation does not seem to be significantly more difficult to solve as compared to the other variants. As the results in Table 8 show, FF was not able to optimally solve any of the 51-node symmetric and 65-node asymmetric instances, and failed to provide feasible solutions for six. This can be explained by the increased complexity of the problem and the size of the corresponding formulation. On the other hand, all instances with 24 nodes, two instances with 51 nodes, and all asymmetric instances with 36 or 65 nodes are solved to optimality by either IP1 or IP2 within the time limit of three hours. Similar to the previous case, although `eil51c.3` and `eil51c.4` have the same number and type of routes, the latter has proved easier to solve compared to the former, despite having more restrictive bounds.

5 Conclusions

This paper described a polynomial transformation of the BRP, a unit-demand vehicle routing problem with lower and upper bounds on the number of nodes visited on each route, first into a GTSP, then into a TSP. The resulting TSP was formulated as an integer program in the space of the natural binary variables only, and one which did not require the usual subtour elimination constraints, if supplemented by side constraints that are written for each cluster of the corresponding GTSP. The procedure described in the paper is such that the BRP amounts to no more than solving a constrained assignment problem with a polynomial number of constraints. The sparsity of the transformation graph also allows one to significantly reduce the number of binary variables in the formulation (typically between 96–98%) which facilitates the solution of the resulting formulation using off-the-shelf solvers.

In this paper, we have shown through computational experimentation that the proposed transformation can be applied to the BRP symmetric and asymmetric costs, single or multiple depots, or homogeneous or heterogeneous fleet of vehicles, although the results presented in this paper suggest that the approach works particularly and consistently well in the following two cases:

- The heterogeneous BRP: In this case, the integer programming model obtained as a result of the transformation can solve more instances to optimality than the flow-based formulations tested in this paper, and in much less computational time. In the heterogeneous case, this means being able to optimally solve symmetric instances of up to 51 nodes and asymmetric instances of up to 65 nodes, either with single or multiple depots.
- Better balanced BRP instances: In this case, the efficiency of the model obtained as a result of the transformation improves as the interval $[Q_{min}, Q_{max}]$ gets tighter, which means the resulting instance is better balanced. This is in contrast to the performance of the flow-based formulations, which become increasingly more difficult to solve as the same interval is made narrower.

Acknowledgments

Thanks are due to the Associate Editor and to two anonymous reviewers who provided comments on an earlier version of this paper. The authors gratefully acknowledge the Faculty Strategic Research Fund provided by the University of Southampton in carrying out this research.

References

- Baldacci, R., M. Battarra, D. Vigo. 2008. Routing a heterogeneous fleet of vehicles. B. Golden, S. Raghavan, E. Wasil, eds., *The vehicle routing problem: latest advances and new challenges*. Operations Research/Computer Science Interfaces, Springer, New York, 3–27.
- Behzad, A., M. Modarres. 2002. A new efficient transformation of generalized traveling salesman problem into traveling salesman problem. *Proceedings of the 15th International Conference of Systems Engineering*. ICSE, Springer, Las Vegas, 43–46.
- Bektaş, T. 2006. The multiple traveling salesman problem: an overview of formulations and solution procedures. *Omega* **34** 209–219.
- Bektaş, T. 2012. Formulations and Benders decomposition algorithms for multidepot salesmen problems with load balancing. *European Journal of Operational Research* **216** 83–93.
- Bellmore, M., S. Hong. 1977. A note on the symmetric multiple travelling salesman problem with fixed charges. *Operations Research* **25** 871–874.
- Benavent, E., A. Martinez-Sykora. 2007. Problema del viajante multiple con multiples depositos. Master's thesis, University of Valencia, Spain.
- Calvete, Herminia I, Carmen Gale, Maria-Jose Oliveros, Belén Sánchez-Valverde. 2007. A goal programming approach to vehicle routing problems with soft time windows. *European Journal of Operational Research* **177**(3) 1720–1733.
- Cordeau, J.-F., G. Laporte, M.W.P. Savelsbergh, D. Vigo. 2007. Vehicle routing. C. Barnhart, G. Laporte, eds., *Transportation, Handbooks in Operations Research and Management Science*. Elsevier, Amsterdam, 367–428.
- Dimitrijević, V., Z. Šarić. 1997. An efficient transformation of the generalized traveling salesman problem into the traveling salesman problem on digraphs. *Information Sciences* **102** 105–110.
- Gouveia, L., J. Riera-Ledesma, J.J. Salazar-González. 2013. Reverse multistar inequalities and vehicle routing problems with a lower bound on the number of customer per route. *Networks* **61** 309–321.
- Gouveia, L., J.-J. Salazar-González. 2013. Polynomial-time separation of enhanced reverse multistar inequalities. *Operations Research Letters* **41**(3) 294–297.
- Groër, C., B. Golden, E. Wasil. 2009. The balanced billing cycle vehicle routing problem. *Networks* **54**(4) 243–254.
- GuoXing, Y. 1995. Transformation of multidepot multisalesmen problem to the standard travelling salesman problem. *European Journal of Operational Research* **81** 557–560.

- Hong, S., M.W. Padberg. 1977. A note on the symmetric multiple traveling salesman problem with fixed charges. *Operations Research* **25** 871–874.
- Jonker, R., T. Volgenant. 1988. An improved transformation of the symmetric multiple traveling salesman problem. *Operations Research* **36** 163–167.
- Kritikos, Manolis N, George Ioannou. 2010. The balanced cargo vehicle routing problem with time windows. *International Journal of Production Economics* **123**(1) 42–51.
- Laporte, G. 1992a. The traveling salesman problem: An overview of exact and approximate algorithms. *European Journal of Operational Research* **59**(2) 231–247.
- Laporte, G. 1992b. The vehicle routing problem: An overview of exact and approximate algorithms. *European Journal of Operational Research* **59**(3) 345–358.
- Laporte, G. 2007. What you should know about the vehicle routing problem. *Naval Research Logistics* **54**(8) 811–819.
- Laporte, G., F. Semet. 1999. Computational evaluation of a transformation procedure for the symmetric generalized traveling salesman problem. *INFOR* **37**(2) 114–120.
- Noon, C.E., J.C. Bean. 1993. An efficient transformation of the generalized traveling salesman problem. *INFOR* **31**(1) 39–44.
- Okonjo-Adigwe, C. 1988. An effective method of balancing the workload amongst salesmen. *Omega* **16**(2) 159–163.
- Rao, M.R. 1980. A note on multiple travelling salesmen problem. *Operations Research* **28** 628–632.
- Toth, P., D. Vigo. 2002. *The vehicle routing problem*. SIAM Monographs on Discrete Mathematics and Applications, Philadelphia.

Appendix

We present formulations for the four variants of the BRP discussed in the main body of text.

Single depot and homogeneous fleet

This formulation is presented in Gouveia et al. (2013). In this case, $D = \{d\}$ (single depot), $Q_{min}^{d,r} = Q_{min}$ and $Q_{max}^{d,r} = Q_{max}$. The formulations uses two sets of variables, namely a binary variable x_{ij} which takes the value 1 if one vehicle travels arc (i, j) and a continuous variable f_{ij} defining unit flows of a commodity from the depot in decreasing units.

The formulation is presented below.

$$\text{Minimize} \quad \sum_{(i,j) \in A} c_{ij} x_{ij} \quad (5.1)$$

$$\text{subject to} \quad \sum_{j \in C \cup D} x_{ij} = 1, \quad i \in C, \quad (5.2)$$

$$\sum_{j \in C \cup D} x_{ji} = 1, \quad i \in C, \quad (5.3)$$

$$\sum_{j \in C \cup D} x_{dj} = m_d, \quad (5.4)$$

$$\sum_{j \in C \cup D} x_{jd} = m_d, \quad (5.5)$$

$$\sum_{j \in C \setminus \{i\} \cup D} f_{ij} - \sum_{j \in C \setminus \{i\} \cup D} f_{ji} = 1, \quad i \in C, \quad (5.6)$$

$$Q_{min} x_{1i} \leq f_{1j} \leq Q_{max} x_{1i}, \quad i \in C, \quad (5.7)$$

$$x_{ij} \leq f_{ij} \leq (Q_{max} - 1)x_{ij}, \quad i, j \in C. \quad (5.8)$$

Multiple depots with homogeneous fleet

This model is the multi-commodity formulation presented in Bektaş (2012). In this case $Q_{min}^{d,r} = Q_{min}^{d,t} (= Q_{min})$ and $Q_{max}^{d,r} = Q_{max}^{d,t} (= Q_{max}), 1 \leq r < t \leq m_d$. The formulation is defined on a slightly modified graph $G = (D, C, D', A)$ in which a copy for each depot is added. The set D' denotes the set of copies of the depots in D . For each $i \in D$, we denote by $r_i \in D'$ the corresponding copy. The set of arcs is $A = \{(i, j) | i \in D \cup C, j \in C\} \cup \{(i, j) | i \in C, j \in D'\}$. The set of commodities is shown by $Q = \{q_1, \dots, q_{|D|}\}$. This formulation uses a binary variable y_{ij} which takes the value 1 if one vehicle travels arc (i, j) , and 0 otherwise, a continuous variable $x_{ij}^{q_d}$ denotes the amount of flow of commodity q_d on arc (i, j) and a continuous variable $f_{ij}^{q_d}$ describes the flow of commodities to each customer.

$$\text{Minimize} \quad \sum_{(i,j) \in A} c_{ij} y_{ij} \quad (5.9)$$

$$\text{subject to} \quad \sum_{j \in C \cup D'} x_{ij}^{q_d} - \sum_{j \in C \cup D} x_{ji}^{q_d} = \begin{cases} m_d, & \text{if } i = d \\ -m_d, & \text{if } i = r_d \\ 0, & \text{otherwise} \end{cases}, \quad i \in C, q_d \in Q, \quad (5.10)$$

$$x_{ij}^{q_d} \leq y_{ij}, \quad i \in C, q_d \in Q \quad (5.11)$$

$$\sum_{j \in C \cup D'} y_{ij} = 1, \quad i \in C \quad (5.12)$$

$$\sum_{j \in C \cup D} y_{ji} = 1, \quad i \in C \quad (5.13)$$

$$\sum_{j \in C} y_{ij} = m_i, \quad i \in D \quad (5.14)$$

$$\sum_{j \in C} y_{ji} = m_i, \quad i \in D' \quad (5.15)$$

$$\sum_{j \in C \cup D} f_{ji}^i - \sum_{j \in C \cup D'} f_{ij}^i = 1, \quad i \in C, \quad (5.16)$$

$$\sum_{k \in C \cup D} f_{kj}^i - \sum_{k \in C \cup D'} f_{jk}^i = 0, \quad i, j \in C, \quad (5.17)$$

$$\sum_{k \in D, j \in C} f_{kj}^i = 1, \quad i \in C, \quad (5.18)$$

$$f_{ij}^k \leq y_{ij}, \quad (i, j) \in A, k \in C, \quad (5.19)$$

$$\sum_{i \in C} f_{kj}^i \geq Q_{min} y_{kj}, \quad k \in D, j \in C \quad (5.20)$$

$$\sum_{k \in C, k \neq i} f_{ij}^k \leq (Q_{max} - 1)y_{ij}, \quad i \in C, j \in C \cup D'. \quad (5.21)$$

Single depot with heterogeneous fleet

A formulation in Baldacci et al. (2008) can be found for the CVRP with heterogeneous fleet, of which we will consider the unit demand version augmented with the bounds $Q_{min}^{d,r}$ on the number of customers visited by each vehicle. The

formulation below is therefore a combination of those presented in Baldacci et al. (2008) and Gouveia and Salazar-González (2013). In this case, $D = \{d\}$ (single depot). This formulation uses a binary variable x_{ij} equal to 1 if vehicle k travels arc (i, j) , and 0 otherwise, and a continuous variable f_{ij} defining the flow of a single commodity.

$$\text{Minimize} \quad \sum_{(i,j) \in A} \sum_{k=1}^{m_d} c_{ij} x_{ij}^k \quad (5.22)$$

$$\text{subject to} \quad \sum_{k=1}^{m_d} \sum_{j \in C \cup D} x_{ij}^k = 1, \quad i \in C, \quad (5.23)$$

$$\sum_{k=1}^{m_d} \sum_{j \in C \cup D} x_{ji}^k = 1, \quad i \in C, \quad (5.24)$$

$$\sum_{k=1}^{m_d} \sum_{j \in C \cup D} x_{dj}^k = m_d, \quad (5.25)$$

$$\sum_{k=1}^{m_d} \sum_{j \in C \cup D} x_{jd}^k = m_d, \quad (5.26)$$

$$\sum_{j \in C \cup D} x_{ij}^k = \sum_{j \in C \cup D} x_{ji}^k, \quad i \in C, k \in \{1, \dots, m_d\}, \quad (5.27)$$

$$\sum_{j \in C \setminus \{i\} \cup D} f_{ij} - \sum_{j \in C \setminus \{i\} \cup D} f_{ji} = 1, \quad i \in C, \quad (5.28)$$

$$Q_{min} \sum_{k=1}^{m_d} x_{1i}^k \leq f_{1j} \leq Q_{max} \sum_{k=1}^{m_d} x_{1i}^k, \quad i \in C, \quad (5.29)$$

$$\sum_{k=1}^{m_d} x_{ij}^k \leq f_{ij} \leq (Q_{max} - 1) \sum_{k=1}^{m_d} x_{ij}^k, \quad i, j \in C. \quad (5.30)$$

$$(5.31)$$

Multiple depots with heterogeneous fleet

This formulation a combination of the multi-commodity formulation presented in Bektaş (2012) and that by Baldacci et al. (2008) for heterogeneous vehicle routing. The graph is again augmented as $G = (D, C, D', A)$, in which a copy for each depot $i \in D$ appears $r_i \in D'$. The set of arcs is $A = \{(i, j) | i \in D \cup C, j \in C\} \cup \{(i, j) | i \in C, j \in D'\}$. The set of commodities is shown by $Q = \{q_1, \dots, q_{|D|}\}$. In this formulation, a binary variable $y_{ij}^{k,d}$ takes the value 1 if vehicle k of depot d travels arc (i, j) , and 0 otherwise, a continuous variable $x_{ij}^{q,d}$ denotes the amount of flow of commodity q_d on arc (i, j) , and a continuous variable $f_{ij}^{q,d}$ describing commodity flows.

$$\text{Minimize} \quad \sum_{d=1}^{|D|} \sum_{k=1}^{m_d} \sum_{(i,j) \in A} c_{ij} y_{ij}^{k,d} \quad (5.32)$$

$$\text{subject to} \quad \sum_{j \in C \cup D'} x_{ij}^{q,d} - \sum_{j \in C \cup V_D} x_{ji}^{q,d} = \begin{cases} m_d, & \text{if } i = d \\ -m_d, & \text{if } i = r_d \\ 0, & \text{otherwise} \end{cases}, \quad i \in C, q_d \in Q, \quad (5.33)$$

$$x_{ij}^{q,d} \leq y_{ij}, \quad i \in C, q_d \in Q \quad (5.34)$$

$$\sum_{d=1}^{|D|} \sum_{k=1}^{m_d} \sum_{j \in C \cup D'} y_{ij}^{k,d} = 1, \quad i \in C \quad (5.35)$$

$$\sum_{d=1}^{|D|} \sum_{k=1}^{m_d} \sum_{j \in C \cup V_D} y_{ji}^{k,d} = 1, \quad i \in C \quad (5.36)$$

$$\sum_{k=1}^{m_i} \sum_{j \in C} y_{ij}^k = m_i, \quad i \in D \quad (5.37)$$

$$\sum_{k=1}^{m_i} \sum_{j \in C} y_{ji}^k = m_i, \quad i \in D' \quad (5.38)$$

$$\sum_{j \in C \cup V_D} y_{ji}^{k,d} = \sum_{j \in C \cup D'} y_{ij}^{k,d}, \quad i \in C, d \in D, k \in \{1, \dots, m_d\} \quad (5.39)$$

$$\sum_{j \in C \cup D} f_{ji}^i - \sum_{j \in C \cup D'} f_{ij}^i = 1, \quad i \in C, \quad (5.40)$$

$$\sum_{k \in C \cup D} f_{kj}^i - \sum_{k \in C \cup D'} f_{jk}^i = 0, \quad i, j \in C, \quad (5.41)$$

$$\sum_{k \in D, j \in C} f_{kj}^i = 1, \quad i \in C, \quad (5.42)$$

$$f_{ij}^t \leq \sum_{d=1}^{|D|} \sum_{k=1}^{m_d} y_{ij}^{k,d}, \quad (i, j) \in A, t \in C, \quad (5.43)$$

$$\sum_{i \in C} f_{tj}^i \geq Q_{min} \sum_{d=1}^{|D|} \sum_{k=1}^{m_d} y_{ij}^{k,d}, \quad t \in D, j \in C \quad (5.44)$$

$$\sum_{k \in C, k \neq i \neq j} f_{ij}^k \leq (Q_{max} - 1) \sum_{d=1}^{|D|} \sum_{k=1}^{m_d} y_{ij}^{k,d}, \quad i \in C, j \in C \cup D'. \quad (5.45)$$