Ann Oper Res (2016) 246:101–126
DOI 10.1007/s10479-014-1682-7

A Lagrangian relaxation approach for expansion of a highway network

Eusebio Angulo · Ricardo García-Ródenas · José Luis Espinosa-Aranda

Published online: 20 August 2014

Abstract This paper deals with the problem of improving an existing road network in the context of strategic planning through the creation of new highway corridors. To address this problem we analyse three mixed integer programming models. The so-called [P1] is the classical capacitated multicommodity network design model. The model named [P2] imposes on [P1] the location of a single (main path) highway corridor in the road network and [P3] adds to [P2] a set of sub-tour breaking constraints. The stated goal is to minimize the total travel time for a known origin-destination demand matrix with a given budget. In this paper we propose an efficient method for [P3], based on a Lagrangian relaxation, to obtain easily-solved sub-problems. A cutting-plane method for solving the Lagrangian sub-problems is proposed. This method generates valid cuts until an optimal solution is found. The Lagrangian dual problem is solved using the sub-gradient optimization method. A case study has been carried out for the region of Castilla-La Mancha (Spain). Computational comparisons between the proposed method and a state-of-the-art mixed-integer code are presented. The Lagrangian relaxation approach is found to be capable of generating good feasible solutions to the case study within a reasonable computational time.

Keywords Highway corridor location · Multicommodity network design · Network expansion · Lagrangian relaxation · Demand covering

E. Angulo (✉) · R. García-Ródenas · J. L. Espinosa-Aranda
Escuela Superior de Informática, Departamento de Matemáticas, Universidad de Castilla La Mancha, 13071 Ciudad Real, Spain
e-mail: Eusebio.Angulo@uclm.es

R. García-Ródenas
e-mail: Ricardo.Garcia@uclm.es

J. L. Espinosa-Aranda
e-mail: JoseL.Espinosa@uclm.es

1 Introduction

The problem of highway corridor location requires a huge amount of resources and consequently the new infrastructures are expected to provide the best possible service for the spatial demand distribution.

This problem is addressed in the literature with microscopic highway network planning and is known as the highway alignment optimization (HAO) problem (see Jong 1998). HAO is a microscopic highway route design problem whose input data are the highway endpoints and costs involved. The goals and constraints involved in HAO problems are analysed in detail in Jong (1998) and Kang (2008). The majority of studies dealing with HAO problems use Genetic Algorithms in the geographic information system (GIS) environment (see Jha and Schonfeld 2000; Jha et al. 2001; Jha 2001). These models assume the creation of the network from scratch.

Recently, Angulo et al. (2012, 2014) apply continuous location models (see Mesa and Boffey 1996) for solving HAO. The highway endpoints are not inputs in Angulo et al. (2012), and the captured demand is taken into account. In Angulo et al. (2014), the problem of locating new highway corridors simultaneously with an existing rural highway network is addressed by means of a bi-level model.

Therefore, the continuous location models assume many simplifications in user route-choice modelling. The so-called *Network Design Problem* (NDP) (see Yang and Bell 1998) overcomes this weakness and locates a set of facilities in the links of a road network taking into account the route decisions of users. Solanki et al. (1998) apply an uncongested NDP to the choice of road links and timing. NDP locates a set of road links, possibly disconnected and not completely defined as a highway to be built. A proper approach should consist of locating an adequate path on a network that minimizes the network travel cost (or travel time). Practical examples of this problem can be found in transportation literature as *Rapid Transit Network Design (RTND)* (see Laporte et al. 2000, 2002, 2005, 2007; Escudero and Muñoz 2009; Marín and García-Ródenas 2009; Gutíerrez-Jarpa et al. 2010; Laporte et al. 2011a, b) and telecommunications networks (see Erdemir et al. 2008; Konak 2012 among others).

This paper deals with the problem of expanding an existing uncongested rural highway network in a strategic planning context through the creation of new stretches of highway corridors. Our problem is included in the category of covering path models. These problems determine routes on a graph so that the demand associated is (node or trip) covered or, at least, maximized. The literature is rich in these problems (see Current and Marsh 1993), and the two main types proposed are:

– *The maximum covering-shortest path problem (MCSPP)* was first introduced by Current et al. (1985). In this problem, a path is located between an origin and a destination with a bicriterion approach. The first objective is to minimize the construction cost of the links of the path, and the second is the maximizing of node demand satisfaction. Examples of MCSPP can be found in Current et al. (1984), Current et al. (1994), Boffey and Narula (1998).
– *The Hierarchical Network Design Problem (HNDP)* consists of simultaneously locating a minimum cost path visiting two or more nodes and a minimal-cost forest (one or more trees) connecting all remaining nodes to the path. HNDPs are proposed in Current et al. (1986), Duin and Volgenant (1989), Pirkul et al. (1991) and Sancho (1997).

The MCSPP and HNDP models may be inappropriate to be used to study the problem dealt with in this paper because they do not allow the selection by users of combined routes to be considered, i.e. using partly the new located highway and partly the existing network.

However they share an essential characteristic; the solutions split into a main path and a number of sub-tours. As the objective is to locate a highway that extends the existing road network, these sub-tours must be removed. This task adds complexity (see Timothy 2013) to the above problems.

Our problem involves location and network design decisions which simultaneously consider the location of facilities and the design of the underlying network to minimize the total travel time. Contreras and Fernández (2012) presents a unified framework for the general network design problem involving combined location and network design decisions.

The above covering path models are of great computational complexity ($\mathcal{NP}$-hard) which requires the use of heuristic methods (for example Dufourd et al. 1996; Solanki et al. 1998; Crainic et al. 2000; Bruno et al. 2002). Laporte et al. (2000, 2011b) provide a good review of methods used to solve these models. Recently, Fernández and Marín (2003) addresses the location of a bus line with multisource demand, which subsumes an important class of strategic path location models. The proposed method is heuristic, especially designed for acyclic networks. Their examples have a size of up to 200 nodes, and the gaps between the best solutions and the bounds are rather large (between 3.7 and 11.6 %).

Exact methods are scarcely shown in the literature and the numerical experiments described are carried out only in very small-scale networks which comprise less than a dozen nodes, as can be seen Bruno et al. (1998), Laporte et al. (2007), Marín (2007) and Marín et al. (2009). Recently, Gutíerrez-Jarpa et al. (2010) addresses the problem of locating a free path (with no pre-set extreme nodes) on a network, in such a way as to minimize the cost and maximize the traffic captured by the path. These authors, in order to avoid these sub-tours, add a sub-tour breaking constraint. This constraint sends a unit of flow from an arbitrary origin to each of the nodes that are visited by the path.

In this paper three mixed integer programming (MIP) problems have been analysed, [P1], [P2] and [P3]. The so-called [P1] model is the capacitated multicommodity NDP. In this paper, it is assumed that the level of congestion is negligible (interurban highways). The stated goal is to aid investment decisions in new infrastructures, minimizing the total travel time for a fixed budget and a known origin-destination demand matrix. The solution of [P1] is a set of (disconnected) links which represent the new set of highways. The [P2] model locates a main path plus a set of disjointed sub-tours, and the [P3] model adds to the [P2] model a set of constraints to avoid the sub-tours.

Taking into account the high computational cost for solving this type of model, a method for reducing the computational burden is required. The Lagrangian relaxation approaches have been successfully applied to NDP (see Holmberg and Yuan 2000; Crainic et al. 2001) and location problems (see Contreras et al. 2009; Marín and Pelegrín 1999) for this task. In this paper, a Lagrangian relaxation approach with sub-gradient optimization has been applied to [P3]. The relaxed primal problems are decomposed into two sub-problems: (1) a shortest path problem and (2) a problem consisting of locating a main path on a graph. Sub-problem (2) is solved via a study of valid inequalities (or cuts) of the mixed-integer formulation which break all sub-tours for a given path. The resulting approach is an exact cutting-plane method.

A case study has been carried out for the region of Castilla-La Mancha (Spain), considering 6, 67, 290 population centres with more than 50,000, 5,000 and 1,000 inhabitants respectively. The numerical results show that this problem can be addressed with the proposed method in a reasonable computational time.

In synthesis, the main contributions of this paper consists of formulating [P3] and utilizing an exact method based on Lagrangian relaxation for finding the solution to [P3] for reasonably sized networks in which the state-of-art codes are unsuccessful.

The rest of the paper is organized in the following way: Sect. 2 sets out in detail the network design models constructed in order to approach the problem. The Lagrangian relaxation approach for [P3] is shown in Sect. 3, to obtain a main path highway corridor in the network. In Sect. 4, computational experiments are carried out on a real case study and on a synthetic network. Finally, the last section of the paper provides some conclusions.

2 Network design models

2.1 The problem

Suppose we wish to expand the highway network in a given geographical area. Within this region a subset of population centres (cities) is considered, distributed on the plane. These points define the set of nodes $\mathcal{N}$. An origin-destination demand matrix is assumed to exist between the set of nodes. Each origin-destination pair is denoted by $\omega = (o, d)$ and $\mathcal{W}$ is the set of origin-destination pairs. Let g_ω be the demand at *O–D* pair w.

The existing network is modelled by a graph $\mathcal{G}_A = (\mathcal{N}, \mathcal{A})$. The node set coincides with the set of population centres $\mathcal{N}$ served by the existing road network. The directed link $a = (i, j)$ will be included in the set of links, i.e $a \in \mathcal{A}$ if there exists a road which connects towns i and j without going through any other population centre. Note that if cities i and j are joined by a highway the two directed links (i, j) and (j, i) are included in $\mathcal{A}$. Each link is classified according to a typology which determines the average journey speed in that link, and, as a result, the journey time c_a at link $a \in \mathcal{A}$ is calculated as a function of distance and link type.

Let $\hat{\mathcal{B}}$ be the set of new links which can be constructed, which will be determined "a priori" by the user. In this paper we consider all the links of $\mathcal{N} \times \mathcal{N}$ whose Euclidean distance is less than a given constant. A link $b = \{i, j\} \in \hat{\mathcal{B}}$ means that a new highway between two centres i and j can be built. The routing of the demand forces us to consider two directed links associated with b, which are (i, j) and (j, i).

Abusing the notation we also denote by $b = (i, j)$, with $b' = (j, i)$ and $\mathcal{B}$ the set of all these directed links. The typology of these new links and the corresponding travel times c_a with $a \in \mathcal{B}$ are assumed to be known. The problem is to find a highway corridor in the $\mathcal{G}_B$ network that meets a given budget $B_0 > 0$ and optimizes the service of the network $\mathcal{G}_A \cup \mathcal{G}_B$. For the sake of clarity, the notation used in the formulation of the network design model is set out in Table 1.

2.2 [P1]: Corridor identification using a multicommodity network design model

In most transportation applications each origin-destination pair $\omega \in \mathcal{W}$ is associated with a commodity. In our model we consider that the commodities are defined as the demands with the same destination d (analogously this can also be defined with respect to the origins). This approach significantly reduces the number of variables and constraints in the network design formulation.

Let $\mathcal{P}$ be the set of different commodities considered in the network, consisting of all different possible destinations, and let $d \in \mathcal{P}$ be each of these destinations.

To be able to define the flow conservation equations we need to define the total demand at node i travelling to destination d:

Table 1 Notation used in the models [P1], [P2], [P2′] and [P3]

Sets and indexes

$\mathcal{N}$: Set of nodes in the network

$\mathcal{P}$: Set of commodities in the network

$\mathcal{A}$: Set of directed links which represent the existing road network

$\widehat{\mathcal{B}}$: Set of edges which represent the highway corridors capable of being built

$\mathcal{B}$: Set of directed links derived from $\widehat{\mathcal{B}}$

$\mathcal{C}_i^-: = \{c \in \mathcal{A} \cup \mathcal{B} : c = (i, j)\}$. Set of links with origin in node i

$\mathcal{C}_i^+: = \{c \in \mathcal{A} \cup \mathcal{B} : c = (j, i)\}$. Set of links with destination in node i

$\mathcal{N}_{\widehat{\mathcal{B}}}: = \{i \in \mathcal{N} \mid i$ is an endpoint of $b \in \widehat{\mathcal{B}}$, i.e. $b = (i, j) \in \mathcal{B}$ *or* $b = (j, i) \in \mathcal{B}\}$

$\widehat{\mathcal{B}}_i: = \{b \in \widehat{\mathcal{B}} \mid b$ is incident to $i \in \mathcal{N}$, i.e. $b = \{i, j\} \in \widehat{\mathcal{B}}\}$

a: A link of $\mathcal{A} \cup \mathcal{B}$

b: An edge of $\widehat{\mathcal{B}}$ or a directed link of $\mathcal{B}$

b': If $b = (i, j)$ then $b' = (j, i)$

d: A commodity of $\mathcal{P}$

ω: Demand pair between two nodes of the network $\mathcal{G}_{\mathcal{A}}$. The first element is the origin of the demand and the second element the destination

$\mathcal{W}$: Set of all the origin-destination pairs ω

$\mathcal{W}_d$: Set of all the pairs whose destination is node d

Data

c_a: Travel time in link $a \in \mathcal{A} \cup \mathcal{B}$

M: Capacity of link $b \in \widehat{\mathcal{B}}$

g_ω: Demand of pair ω

$\{g_\omega\}_{\omega \in \mathcal{W}}$: The origin-destination (O–D) demand matrix

B_0: Budget for construction of new highway corridors

f_b: Construction cost of link $b \in \widehat{\mathcal{B}}$

Variables

x_a^d: Number of users of commodity d in link $a \in \mathcal{A} \cup \mathcal{B}$

y_b: Has value 1 if a new highway corridor is built in link $b \in \hat{\mathcal{B}}$ and 0 otherwise

δ_i: Degree of node i in the new road transport network, that is, the number of edges of $\widehat{\mathcal{B}}$ incident to the node i

γ_i: Shows whether node i is added to the new road transport network. This variable must satisfy: If $\delta_i > 0 \Rightarrow \gamma_i = 1$

$$g_i^d = \begin{cases} 0 & \text{if the node } i \text{ is neither origin nor destination} \\ -g_\omega & \text{if } i \text{ is the origin of pair } \omega\text{, i.e.} \\ \sum_{\omega \in \mathcal{W}_d} g_\omega & \text{if } i = d \end{cases} \tag{1}$$

where $\mathcal{W}_d$ is the set of all origin-destination pairs whose destination is node d.

In the model there are two types of variables: (a) the location variables y_b, which determine whether a certain stretch of highway will or will not be built, and (b) routing variables x_a^d, which determine the number of users of commodity d in link $a \in \mathcal{A} \cup \mathcal{B}$.

The proposed model **[P1]** has the following formulation:

$$Minimize \quad Z = \sum_{d \in \mathcal{P}} \sum_{a \in \mathcal{A} \cup \mathcal{B}} c_a x_a^d \tag{2}$$

$$subject\ to: \quad \sum_{d \in \mathcal{P}} (x_b^d + x_{b'}^d) \leq M y_b, \quad b \in \widehat{\mathcal{B}}, \quad d \in \mathcal{P} \tag{3}$$

$$\sum_{c \in \mathcal{C}_i^+} x_c^d - \sum_{c \in \mathcal{C}_i^-} x_c^d = g_i^d, \quad i \in \mathcal{N}, \quad d \in \mathcal{P} \tag{4}$$

$$x_a^d \geq 0, \quad a \in \mathcal{A} \cup \mathcal{B}, \quad d \in \mathcal{P} \tag{5}$$

$$\sum_{b \in \widehat{\mathcal{B}}} y_b f_b \leq B_0 \tag{6}$$

$$y_b \in \{0, 1\}, \quad b \in \widehat{\mathcal{B}}. \tag{7}$$

The objective function (2) represents the network design criterion. This criterion consists in minimizing the total time in the network to satisfy the given O–D demand matrix $\{g_\omega\}_{\omega \in \mathcal{W}}$. This forces the model to implicitly calculate the shortest paths in the resulting network. The aim of constraint (3) is twofold: (a) it shows that if a section of highway corridor is not built then this link is not available to the users, and (b) if the link is built it establishes a capacity M for this arc. Constraint (4) expresses the requirements of flow conservation and demand satisfaction and guarantees that users who leave an origin reach their destination. Constraint (6) is the budgetary constraint and it indicates that investment in the network cannot exceed the given budget B_0. Constraints (5) and (7) represent the type of variables in the problem. The result is a large scale mixed linear programming model.

To simplify we define:

$$\mathcal{H} := \{(x, y) : (x, y) \text{ which satisfy the restrictions (3)–(7)}\}.$$

2.3 [P2]: Location of a single main path and sub-tours in a graph

The solution to model [P1] identifies several corridors which are in need of improvement. In some highway location problems the solution is required to be polygonal. This polygonal (main path) will mark a preliminary alignment according to demand, which should be refined in the next stages of the planning process.

In order to adapt model [P1] to this situation so that it will choose a single main path in the form of a polygonal, we introduce the following sets and variables:

$$\mathcal{N}_{\widehat{\mathcal{B}}} = \left\{ i \in \mathcal{N} : \text{If } i \text{ is an endpoint of } b \in \widehat{\mathcal{B}} \right\}$$
$$\widehat{\mathcal{B}}_i = \left\{ b \in \widehat{\mathcal{B}} : \text{If } b \text{ is incident to } i \in \mathcal{N}_{\widehat{\mathcal{B}}} \right\}$$

$\delta_i \equiv$ is the degree of node i, that is, the number of edges of $\widehat{\mathcal{B}}$ that are incident with node i.
$\gamma_i \equiv$ shows whether node i is in the new network. This variable must satisfy:

$$\text{If } \delta_i > 0 \Rightarrow \gamma_i = 1 \tag{8}$$

We construct a new model, **[P2]**, introducing a new set of restrictions to [P1]:

$$Minimize \quad Z = \sum_{d \in \mathcal{P}} \sum_{a \in \mathcal{A} \cup \mathcal{B}} c_a x_a^d \tag{9}$$

$$subject\ to: \quad (x, y) \in \mathcal{H} \tag{10}$$

$$\sum_{i \in \mathcal{N}_{\widehat{\mathcal{B}}}} \gamma_i = 1 + \sum_{b \in \widehat{\mathcal{B}}} y_b \tag{11}$$

$$\sum_{b \in \widehat{\mathcal{B}}_i} y_b = \delta_i, \quad i \in \mathcal{N}_{\widehat{\mathcal{B}}} \tag{12}$$

$$\delta_i \leq 2, \quad i \in \mathcal{N}_{\widehat{\mathcal{B}}} \tag{13}$$

$$\gamma_i + \gamma_j \geq 2y_b, \quad b = \{i, j\},\ b \in \widehat{\mathcal{B}} \tag{14}$$

$$\gamma_i \in \{0, 1\}, \quad i \in \mathcal{N}_{\widehat{\mathcal{B}}} \tag{15}$$

$$\delta_i \geq 0, \quad i \in \mathcal{N}_{\widehat{\mathcal{B}}}. \tag{16}$$

Constraint (12) is the definition of degree. Constraint (13) requires that the degree of the vertices of the solution be less than or equal to 2, and so the solutions which satisfy (11) and (13) are a single polygonal and the union of disconnected cycles. Constraints (14) and (15) are constraints necessary to model correctly the variable γ_i defined in (8). Thus, if $\delta_i > 0$, by constraint (12) at least one link $b = \{i, j\}$ must exist such that $y_b = 1$. In that case, applying constraint (14) to this link b would yield $\gamma_i + \gamma_j \geq 2 \cdot 1$, since constraint $\gamma_i, \gamma_j \in \{0, 1\}$ forces that $\gamma_i = \gamma_j = 1$.

Constraints (15) and (16) represent the type of variables in the problem.

The solution (γ, δ, y) defines the sub-graph $\mathcal{G}'_{\widehat{\mathcal{B}}} = (\mathcal{N}', \widehat{\mathcal{B}}')$ of $\mathcal{G}_{\widehat{\mathcal{B}}}$ in the following way: the link $b \in \widehat{\mathcal{B}}$ satisfies $y_b = 1$ if and only if $b \in \widehat{\mathcal{B}}'$. The node $i \in \mathcal{N}'$ if and only if $\gamma_i = 1$ holds.

Theorem 1 *A solution $\mathcal{G}'_{\widehat{\mathcal{B}}}$ of [P2] comprises a main path and the union of disjointed cycles.*

Proof Let us suppose that $\mathcal{G}'_{\widehat{\mathcal{B}}}$ contains k connected components. $(\mathcal{N}^j_{\widehat{\mathcal{B}}}, \widehat{\mathcal{B}}_j)$ denotes for each $j \in \{1, \ldots, k\}$ each connected component of $\mathcal{G}'_{\widehat{\mathcal{B}}}$. Constraint (11) requires that the number of vertices of the solution (γ, δ, y) be equal to the number of links plus one. We shall look at constraint (11) for each of the components of the graph $\mathcal{G}'_{\widehat{\mathcal{B}}}$.

Firstly, as each connected component of the graph $\mathcal{G}'_{\widehat{\mathcal{B}}}$ must contain a tree (as it is connected), the following inequalities are satisfied:

$$\sum_{i \in \mathcal{N}^j_{\widehat{\mathcal{B}}}} \gamma_i \leq 1 + \sum_{b \in \widehat{\mathcal{B}}_j} y_b \qquad j \in \{1, \ldots, k\}. \tag{17}$$

Constraint (17) may be satisfied with equality [case $(a.1)$] or with strict inequality [case $(a.2)$].

$$\text{Case } (a.1) \quad \sum_{i \in \mathcal{N}^j_{\widehat{\mathcal{B}}}} \gamma_i = 1 + \sum_{b \in \widehat{\mathcal{B}}_j} y_b \qquad j \in \{1, \ldots, k\}.$$

$$\text{Case } (a.2) \quad \sum_{i \in \mathcal{N}^j_{\widehat{\mathcal{B}}}} \gamma_i < 1 + \sum_{b \in \widehat{\mathcal{B}}_j} y_b \Rightarrow$$

$$\sum_{i \in \mathcal{N}^j_{\widehat{\mathcal{B}}}} \gamma_i \leq \sum_{b \in \widehat{\mathcal{B}}_j} y_b \qquad j \in \{1, \ldots, k\}.$$

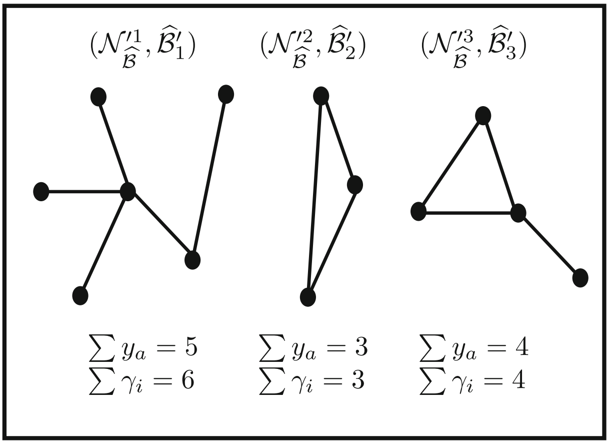


Fig. 1 Example of $\mathcal{G}'_{\widehat{\mathcal{B}}}$ graph applying constraint (11)

Constraint (11) requires that for a single connected component $j' \in \{1, \ldots, k\}$ Case $(a.1)$ holds and the rest of the components j must satisfy:

$$\sum_{i \in \mathcal{N}^{j}_{\widehat{\mathcal{B}}}} \gamma_i = \sum_{a \in \widehat{\mathcal{B}}_j} y_b \qquad j \in \{1, \ldots, k\}, \qquad j \neq j'.$$

Figure 1 shows an example of a $\mathcal{G}'_{\widehat{\mathcal{B}}}$ graph. This graph contains 3 components, of which the first satisfies Case $(a.1)$ and the other two Case $(a.2)$.

Constraint (13) requires that the degree of the vertices of each component k be less than or equal to 2, and so component j' will be a path and the remaining components $j \neq j'$ are trees to which an edge has been added, creating a cycle.

Remark 1 Note that there may exist a node i for which $\delta_i = 0$ and $\gamma_i = 1$ both hold. This may mean that node i is isolated and represents the component j' which satisfies case $(a.1)$. In this case the polygonal (main path) is degenerate and the solution is the union of disjoint cycles (sub-tours).

In the following, we assume that there exists a solution (γ, δ, y) which satisfies constraints (11)–(16) such that the sub-graph $\mathcal{G}'_{\widehat{\mathcal{B}}}$ comprises a main path and the union of a finite number of sub-tours. We define:

$$\Theta := \{(\gamma, \delta, y) : (\gamma, \delta, y) \text{ satisfies the constraints (11)–(16)}\}.$$

Now let us consider an alternative formulation of problem [P2], which we shall call **[P2′]** which avoids the problem described in Remark 1. In this model, let

$$\gamma_i = \begin{cases} 1 & \text{if } \delta_i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

to show whether node i is connected in the new network $\mathcal{G}'_{\widehat{\mathcal{B}}}$.

The model consists of:

$$Minimize \quad Z = \sum_{d \in \mathcal{P}} \sum_{a \in \mathcal{A} \cup \mathcal{B}} c_a x_a^d \tag{19}$$

$$\begin{aligned}
subject\ to: \quad & (x, y) \in \mathcal{H} && (20)\\
& \sum_{i \in \mathcal{N}_{\widehat{\mathcal{B}}}} \gamma_i = 1 + \sum_{b \in \widehat{\mathcal{B}}} y_b && (21)\\
& \sum_{b \in \widehat{\mathcal{B}}_i} y_b = \delta_i, \quad i \in \mathcal{N}_{\widehat{\mathcal{B}}} && (22)\\
& \delta_i \leq 2\gamma_i, \quad i \in \mathcal{N}_{\widehat{\mathcal{B}}} && (23)\\
& \gamma_i \leq \delta_i, \quad i \in \mathcal{N}_{\widehat{\mathcal{B}}} && (24)\\
& \gamma_i \in \{0, 1\}, \quad i \in \mathcal{N}_{\widehat{\mathcal{B}}} && (25)\\
& \delta_i \geq 0, \quad i \in \mathcal{N}_{\widehat{\mathcal{B}}}. && (26)
\end{aligned}$$

Next, we show that constraints (23)–(26) are necessary to model correctly the variable γ_i defined in (18),

$$\gamma_i \leq \delta_i \leq 2\gamma_i.$$

If $\delta_i = 0 \Rightarrow \gamma_i \leq 0 \Rightarrow \gamma_i = 0$. As the variables y_b are binary, if $\delta_i > 0$, by constraint (22), δ_i is an integer, and then $1 \leq \delta_i$, so $1 \leq \delta_i \leq 2\gamma_i$, as $\gamma_i \in \{0, 1\}$, gives that $\gamma_i = 1$ in this case.

Theorem 2 *A solution $\mathcal{G}'_{\widehat{\mathcal{B}}}$ of [P2′] comprises a non-degenerate main path and the union of disjointed cycles.*

Proof Let

$$\Theta' := \{(\gamma, \delta, y) : (\gamma, \delta, y) \text{ satisfies the constraints (21)–(26)}\}.$$

We shall prove the inclusion

$$\Theta' \subset \Theta \tag{27}$$

Constraints (20)–(22) and (25)–(26) are identical to those in the previous model [P2]. Constraint (23) requires that the degree of the vertices of the solution be less than or equal to 2 which implies constraint (13). Constraint (14) holds by relationship (18) and thus the relationship (27) holds. By using Theorem 1, the solutions of Θ' comprise a single path and the union of a finite number of disjointed cycles. Equation (18) forces the paths to be non-degenerate.

2.4 [P3]: Location of a main path

The objective of model [P3] is to locate a complete highway corridor by a main path. Model [P3] adds a set of additional constraints to the model [P2] to avoid sub-tours that can appear in their solutions.

We define sub-tour-breaking constraint in the network $\mathcal{G}_{\widehat{\mathcal{B}}}$ by:

$$\mathcal{C} := \left\{ y : \sum_{b \in \widehat{\mathcal{B}}(S)} y_b \leq |S| - 1 : \forall S \subset \mathcal{N} : 3 \leq |S| \leq |N| - 1 \right\}.$$

where $|S|$ represents the number of nodes of the set S and $\widehat{\mathcal{B}}(S)$ is the subset of links of $\widehat{\mathcal{B}}$ connecting any pair of nodes belonging to the set S. Note that one of these constraints must be written for each subset of the set $\mathcal{N}$ with cardinality greater than 2, which makes the number of constraints grow exponentially with the size of the network.

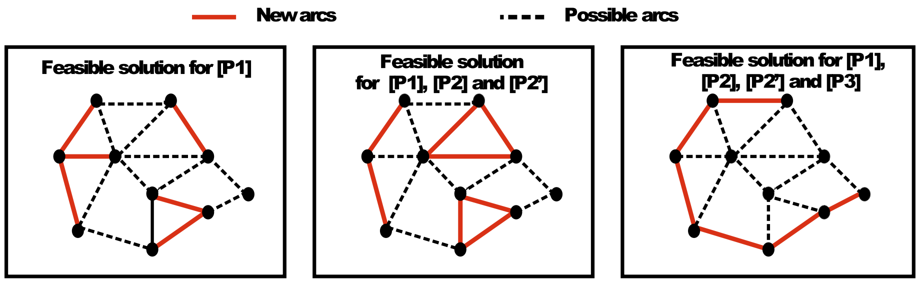


Fig. 2 Examples of the type of feasible solutions for each proposed model

Table 2 Size models [P1], [P2], [P2′] and [P3]

Model	Number of equations	Number of variables	
		Binary	Continuous
[P1]	$1 + \lvert\mathcal{N}\rvert \times \lvert\mathcal{P}\rvert + \lvert\widehat{\mathcal{B}}\rvert$	$\lvert\widehat{\mathcal{B}}\rvert$	$\lvert\mathcal{A} \cup 2\widehat{\mathcal{B}}\rvert \times \lvert\mathcal{P}\rvert$
[P2]	$2 + \lvert\mathcal{N}\rvert \times \lvert\mathcal{P}\rvert + 2\lvert\widehat{\mathcal{B}}\rvert + 2\lvert\mathcal{N}_{\widehat{\mathcal{B}}}\rvert$	$\lvert\widehat{\mathcal{B}}\rvert + \lvert\mathcal{N}_{\widehat{\mathcal{B}}}\rvert$	$\lvert\mathcal{A} \cup 2\widehat{\mathcal{B}}\rvert \times \lvert\mathcal{P}\rvert + \lvert\mathcal{N}_{\widehat{\mathcal{B}}}\rvert$
[P2′]	$2 + \lvert\mathcal{N}\rvert \times \lvert\mathcal{P}\rvert + \lvert\widehat{\mathcal{B}}\rvert + 3\lvert\mathcal{N}_{\widehat{\mathcal{B}}}\rvert$	$\lvert\widehat{\mathcal{B}}\rvert + \lvert\mathcal{N}_{\widehat{\mathcal{B}}}\rvert$	$\lvert\mathcal{A} \cup 2\widehat{\mathcal{B}}\rvert \times \lvert\mathcal{P}\rvert + \lvert\mathcal{N}_{\widehat{\mathcal{B}}}\rvert$
[P3]	$2 + \lvert\mathcal{N}\rvert \times \lvert\mathcal{P}\rvert + 2\lvert\widehat{\mathcal{B}}\rvert + 2\lvert\mathcal{N}_{\widehat{\mathcal{B}}}\rvert + \lvert\mathcal{C}\rvert$	$\lvert\widehat{\mathcal{B}}\rvert + \lvert\mathcal{N}_{\widehat{\mathcal{B}}}\rvert$	$\lvert\mathcal{A} \cup 2\widehat{\mathcal{B}}\rvert \times \lvert\mathcal{P}\rvert + \lvert\mathcal{N}_{\widehat{\mathcal{B}}}\rvert$

The resulting model is as follows:

$$Minimize \quad Z = \sum_{d \in \mathcal{P}} \sum_{a \in \mathcal{A} \cup \mathcal{B}} c_a x_a^d \tag{28}$$

$$subject\ to: \quad (x, y) \in \mathcal{H} \tag{29}$$

$$(\gamma, \delta, y) \in \Theta \tag{30}$$

$$y \in \mathcal{C}, \tag{31}$$

Because the constraints (31) eliminate the sub-tours, in the formulation of [P3] the use of set Θ is equivalent to using set Θ'. It will be seen in Sect. 4.2 that it is computationally more efficient to use Θ than Θ'.

Note that any path is feasible since it does not contain sub-tours and therefore satisfies constraints (29)–(31).

To summarise graphically the above models Fig. 2 depicts three examples of the type of feasible solutions which can be obtained for each of the three proposed models.

Table 2 shows the number of variables and equations for models [P1], [P2], [P2′] and [P3]. Note that the number of variables and equations depends on the number of commodities $|\mathcal{P}|$. Therefore working with origin-destination pairs will have a cost proportional to $|\mathcal{N}|^2$ whereas defining the commodities as the demand to a destination will have a cost proportional to $|\mathcal{N}|$.

3 Lagrangian relaxation approach for model [P3]

Problem [P3] has a suitable form for the use of the Lagrangian relaxation technique. To formulate the problem [P3] in an aggregated way, we denote the flow variable by x and the design variable by y. Note that if the vector y is given the triplet $(\gamma, \delta, y) \in \Theta$ is uniquely determined. For this reason, we omit the variables γ and δ in the formulation of [P3].

We denote

$$\mathcal{X} := \{x : x \text{ satisfies the constraints (4)–(5)}\}.$$
$$\mathcal{Y} := \{y : y \text{ satisfies the constraints (6)–(7), (11)–(16)}\}.$$

and [P3] can be schematically described as:

$$\mathit{Minimize} \quad Z = \sum_{d\in\mathcal{P}} \sum_{a\in\mathcal{A}\cup\mathcal{B}} c_a x_a^d \tag{32}$$

$$\mathit{subject\ to}: \sum_{d\in\mathcal{P}} (x_b^d + x_{b'}^d) \le M y_b, \quad b \in \widehat{\mathcal{B}} \tag{33}$$

$$x \in \mathcal{X} \tag{34}$$

$$y \in \mathcal{Y} \cap \mathcal{C}. \tag{35}$$

Constraint (33) is called *the weak forcing constraint*. Lagrangian relaxation, called *shortest path* (or *flow*) relaxation, is obtained by dualizing the forcing constraint (33). By relaxing the constraint set (33) and using the vector of non-negative multipliers λ associated with it, we obtain the following Lagrangian sub-problem

$$\phi(\lambda) = \mathit{Minimize} \quad Z_x(x,\lambda) - Z_y(y,\lambda) \tag{36}$$

$$\mathit{subject\ to}: \quad x \in \mathcal{X} \tag{37}$$

$$y \in \mathcal{Y} \cap \mathcal{C}. \tag{38}$$

where

$$Z_x(x,\lambda) = \sum_{d\in\mathcal{P}} \sum_{a\in\mathcal{A}} c_a x_a^d + \sum_{d\in\mathcal{P}} \sum_{b\in\widehat{\mathcal{B}}} (\lambda_b + c_b)(x_b^d + x_{b'}^d) \tag{39}$$

$$Z_y(y,\lambda) = \sum_{b\in\widehat{\mathcal{B}}} \lambda_b M y_b \tag{40}$$

The main feature of the Lagrangian sub-problem is its separability into variables x and y. That is

$$\phi(\lambda) = \text{Minimize}_{x\in\mathcal{X}} Z_x(x,\lambda) - \text{Maximize}_{y\in\mathcal{Y}\cap\mathcal{C}} Z_y(y,\lambda) \tag{41}$$

The Lagrangian sub-problem resolves into $|\mathcal{P}|$ shortest path problems and one problem, in y variables only, solvable by a method described in the next subsection. This sub-problem consists of locating a single path (corridor highway) in a graph that expands the existing road network.

The Lagrangian dual related to the relaxation presented can be cast in the form

$$\text{Maximize}_{\lambda\ge 0} \phi(\lambda) \tag{42}$$

For each $\lambda \ge 0$, the value $\phi(\lambda)$ can be obtained by solving the Lagrangian sub-problem; then, one *sub-gradient* $g(\lambda)$ of ϕ in λ can be retrieved from the optimal primal solution (x_λ, y_λ) of the Lagrangian sub-problem. The column vector of constraint mismatches at iteration ℓ constitutes a sub-gradient of the dual function. For the above relaxation, the sub-gradient takes the form

$$g(\lambda) = (\cdots, g_b(\lambda), \cdots)' \in \mathbb{R}^{|\widehat{\mathcal{B}}|} \tag{43}$$

where $'$ is the transpose of a vector and

$$g_b(\lambda) = -M y_{\lambda b} + \sum_{d\in\mathcal{P}} \left[x_{\lambda}{}_b^d + x_{\lambda}{}_{b'}^d \right] \text{ with } b \in \widehat{\mathcal{B}} \tag{44}$$

Table 3 Sub-gradient-based relaxation method for solving [P3]

Step 0	(*Initialization*). Let $\varepsilon^{\text{outer}}$ be a tolerance parameter and $\ell_{max}^{\text{outer}}$ the maximum number of iterations. Set $\ell = 0$. Initialize the dual variables $\lambda_b^\ell = \alpha c_b$ for all $b \in \widehat{\mathcal{B}}$. Set $\phi_L^\ell = -\infty$ and $\phi_U^\ell = +\infty$
Step 1	(*Solution of the Lagrangian sub-problem*). Solve the Lagrangian sub-problem and get the minimizer (x^ℓ, y^ℓ) and the objective function value at the minimizer ϕ^ℓ. Update the lower bound for the objective function of the primal problem $\phi_L^{\ell+1} = \max\{\phi_L^\ell, \phi^\ell\}$
Step 2	(*Upper bound updating*). By the weak duality theorem, one valid upper bound is the value of any primal feasible solution. Let $\psi^\ell := Minimize \quad Z = \sum_{d\in\mathcal{P}} \sum_{a\in\mathcal{A}\cup\mathcal{B}} c_a x_a^d$ $\quad subject\ to \quad : \sum_{d\in\mathcal{P}} (x_b^d + x_{b'}^d) \le M y_b^\ell$ and update $\phi_U^{\ell+1} = \min\{\phi_U^\ell, \psi^\ell\}$ $\quad x \in \mathcal{X}$
Step 3	(*Multiplier updating: Sub-gradient method*). Compute a sub-gradient at λ^ℓ using Eq. (43) $d^{\ell+1} = g(\lambda^\ell) \quad (45)$ The multiplier vector is updated as $\lambda_b^{\ell+1} = \max\{0, \lambda_b^\ell + t_{\ell+1} d_b^{\ell+1}\}$ for all $b \in \widehat{\mathcal{B}} \quad (46)$ where the step size is computed $t_{\ell+1} = \frac{\theta_{\ell+1}(\phi_U^{\ell+1} - \phi_L^{\ell+1})}{\lVert d^{\ell+1}\rVert^2}$ where $\theta_{\ell+1}$ is an iteration-dependent scaling factor. Typically the scalar $\theta_{\ell+1}$ should be assigned a value in the interval (0, 2) in order to ensure convergence
Step 4	Compute the relative gap $RGAP^{\ell+1} = \frac{\phi_U^{\ell+1} - \phi_L^{\ell+1}}{\phi_L^{\ell+1}}$. If $RGAP^{\ell+1} \le \varepsilon^{\text{outer}}$ or $\ell = \ell_{max}^{\text{outer}}$ then Stop, otherwise let $\ell = \ell + 1$ and return to Step 1

In order to solve the Lagrangian dual, we use the well known technique of *sub-gradient optimization*. The sub-gradient algorithm is an iterative method which starts from an initial estimate $\bar{\lambda}$ of the solution and repeats the following five basic steps: (1) select a *tentative ascent direction* d, (2) select a *step size* t, (3) evaluate $\phi(\bar{\lambda} + td)$ and the corresponding sub-gradient; (4) then, move the current point $\bar{\lambda}$ to $\bar{\lambda} + td$; (5) check some stopping criteria.The motivation for using the sub-gradient method is based on its ease of implementation. We have only considered the "basic" sub-gradient method with a standard step size formula (see Polyak 1969). More sophisticated variants to improve the performance are available, but are very sensitive to the parameter adjustment.

The sub-gradient-based relaxation method for solving [P3] is summarized in Table 3.

3.1 A cutting-plane method for solving SUB $_y(\lambda)$

To evaluate the Lagrangian function $\phi(\lambda)$ it is necessary to solve the following sub-problems:

$$\mathbf{SUB}_x(\lambda): \ \text{Minimize}_{x\in\mathcal{X}} Z_x(x, \lambda)$$

and

$$\mathbf{SUB}_y(\lambda): \ \text{Maximize}_{y\in\mathcal{Y}\cap\mathcal{C}} Z_y(y, \lambda)$$

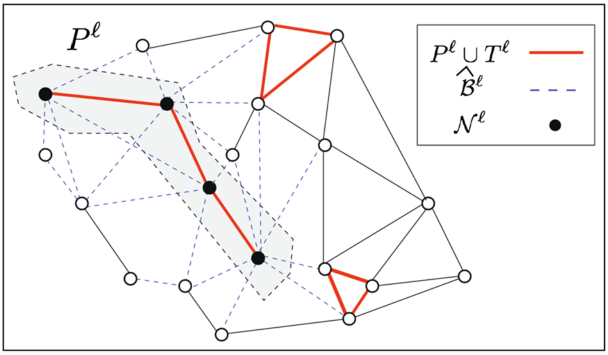


Fig. 3 Illustration of valid cuts

In this section an exact method to solve $\mathrm{SUB}_y(\lambda)$ is proposed. At the beginning, the algorithm relaxes the constraint set $y \in \mathcal{C}$, (possibly with an exponential number of constraints)

$$\mathrm{Maximize}_{y \in \mathcal{Y}} Z_y(y, \lambda) \tag{47}$$

and its solution imposes a valid cut where a new MIP is stated. This procedure is repeated iteratively, generating a sequence of MIPs. In each iteration a new valid cut for problem (47) is added. In a finite number of iterations the algorithm achieves an optimal solution (if it exists) for $\mathrm{SUB}_y(\lambda)$.

Next, the MIP is defined. Let ℓ be the current iteration. We assume that the sub-problem $\mathrm{SUB}_y^\ell(\lambda)$ is given and y^ℓ is an optimal solution. The solution y^ℓ is decomposed as $y^\ell = y^{P\ell} + y^{T\ell}$ where $y^{P\ell}$ is the vector of binary variables which indicates the links associated with the main path and $y^{T\ell}$ is the vector of binary variables associated with the existing sub-tours in the solution y^ℓ. We introduce the following nomenclature

$$\begin{aligned}
P^\ell &:= \{b \in \widehat{\mathcal{B}} : y_b^{P\ell} = 1\}. \text{ The main path.}\\
T^\ell &:= \{b \in \widehat{\mathcal{B}} : y_b^{T\ell} = 1\}. \text{ The set of sub-tours.}\\
\mathcal{N}^\ell &:= \left\{i \in \mathcal{N} : \text{If } i \text{ is visited by the path } P^\ell\right\}\\
\widehat{\mathcal{B}}^\ell &:= \left\{b \in \widehat{\mathcal{B}} : \text{If } b \text{ is incident to some node of } \mathcal{N}^\ell\right\}
\end{aligned}$$

and this is illustrated in Fig. 3.

If $T^\ell \neq \{\emptyset\}$ (see Fig. 3), the solution y^ℓ is not feasible for $\mathrm{SUB}_y(\lambda)$ because it contains sub-tours. We add a valid cut to avoid the solutions where the main path coincides with P^ℓ (in particular the solution y^ℓ). Moreover, the lower bound of $\mathrm{SUB}_y(\lambda)$ is updated as $(Z_y)_L^\ell = \max\{(Z_y)_L^{\ell-1}, Z_y(y^{P\ell})\}$ and we impose that the solutions must be better than the best main path found in the previous iterations. The next sub-problem is stated as:

[$\mathbf{SUB}_y^{\ell+1}(\lambda)$]

$$Maximize\ Z_y(y, \lambda) \tag{48}$$

$$subject\ to:\ y \in \mathcal{Y} \tag{49}$$

Table 4 A cutting-plane (CP) method to solve $\mathrm{SUB}_y(\lambda)$

Step 0	(*Initialization*). Let ε be a tolerance parameter and ℓ_{max} the maximum number of iterations. Set $\ell = 0,\ (Z_y)_L^\ell = -\infty,\ (Z_y)_U^\ell = +\infty$ and $y_\lambda^* = \{\emptyset\}$
Step 1	(*Solution of the* $SUB_y^{\ell+1}(\lambda)$). Solve the sub-problem $\mathrm{SUB}_y^{\ell+1}(\lambda)$. If $\mathrm{SUB}_y^{\ell+1}(\lambda)$ is infeasible then Stop, y_λ^* is an optimal solution for $\mathrm{SUB}_y(\lambda)$, otherwise get the minimizer $y^{\ell+1}$ and the objective function value at the minimizer $\gamma^{\ell+1}$. Decompose $y^{\ell+1} = y^{P,\ell+1} + y^{T,\ell+1}$. Update the upper bound for the objective function of $\mathrm{SUB}_y(\lambda)$ $(Z_y)_U^{\ell+1} = \min\{(Z_y)_U^\ell, \gamma^{\ell+1}\}$
Step 2	(*Updating of relaxed* $SUB_y(\lambda)$. If $y^{T,\ell+1} \neq 0$, compute $P^{\ell+1},\ \widehat{\mathcal{B}}^{\ell+1}$. If $(Z_y)_L^\ell < Z_y(y^{P,\ell+1})$ then $(Z_y)_L^{\ell+1} = Z_y(y^{P,\ell+1})$ and $y_\lambda^* = y^{P,\ell+1}$
Step 3	(*Convergence checking*). Compute the relative gap $$RGAP^{\ell+1} = \frac{(Z_y)_U^{\ell+1} - (Z_y)_L^{\ell+1}}{(Z_y)_L^{\ell+1}}.$$ If $RGAP^{\ell+1} \leq \varepsilon$ or $\ell = \ell_{max}$ then Stop, otherwise let $\ell = \ell + 1$ and return to Step 1

$$\widehat{M} \sum_{b \in \widehat{\mathcal{B}}^s} y_b \geq \sum_{b \in P^s} y_b; \qquad s = 1, \ldots, \ell \tag{50}$$

$$Z_y(y, \lambda) \geq (Z_y)_L^\ell \tag{51}$$

where $\widehat{M}$ is a sufficiently large number. The constraint (50) imposes to find a different path from the previously generated paths.[1] This constraint forces the modification (for example, by expanding it) of the path P^ℓ or locates the main path in the sub-graph associated with the complementary set of nodes to $\mathcal{N}^\ell$.

Theorem 3 *The problem* $SUB_y(\lambda)$ *is infeasible or the CP method achieves an optimal solution for* $SUB_y(\lambda)$ *in a finite number of iterations (letting* $\varepsilon = 0,\ \ell_{max} = +\infty$*)* (Table 4).

Proof We assume that the feasible region of $\mathrm{SUB}_y(\lambda)$ is not empty, that is $\mathcal{Y} \cap \mathcal{C} \neq \{\emptyset\}$. The sub-problem $\mathrm{SUB}_y(\lambda)$ has an optimal solution because $\mathcal{Y} \cap \mathcal{C}$ is a finite set. Let y^* be an optimal solution to $\mathrm{SUB}_y(\lambda)$.

We denote by $\mathcal{Y}^\ell$ the feasible region of the sub-problem $\mathrm{SUB}_y^\ell(\lambda)$. Note that

(i) $\mathcal{Y}^{\ell+1} \subseteq \mathcal{Y}^\ell$ by definition.
(ii) $\mathcal{Y}^0 = \mathcal{Y}$ is a finite set.

As in each iteration the CP method eliminates at least one solution ($y^\ell \notin \mathcal{Y}^{\ell+1}$), the CP method must stop after a finite number of iterations. There must be an integer number $\nu + 1$ such that the sub-problem $\mathrm{SUB}_y^{\nu+1}(\lambda)$ is infeasible.

The solution $y^* \in \mathcal{Y}$ satisfies the constraint (51) because $Z_y(y^{P,\ell_1}, \lambda) = (Z_y)_L^\nu$ for some $\ell_1 \in \{1, \ldots, \nu\}$ and by its optimality (on paths)

$$Z_y(y^*, \lambda) \geq Z_y(y^{P,\ell_1}, \lambda)$$

[1] Note that if a path P^ℓ is not optimal, then all sub-paths of P^ℓ are not optimal because $\lambda \geq 0$.

As $y^* \notin \mathcal{Y}^{\nu+1}$, there exists at least one constraint (50) which is violated. Thus, there exists an integer number $\ell_2 \in \{1, \ldots, \nu\}$ such that

$$\widehat{M} \sum_{b \in \widehat{\mathcal{B}}^{\ell_2}} y_b^* < \sum_{b \in P^{\ell_2}} y_b^*$$

Because the choice of $\widehat{M}$ as a sufficiently large number then $y_b^* = 0$ for all $b \in \widehat{\mathcal{B}}^{\ell_2}$ and $y_{\widetilde{b}}^* = 1$ for some $\widetilde{b} \in P^{\ell_2}$. The solution y^* defines a path and by the definition of $\widehat{\mathcal{B}}^{\ell_2}$ and P^{ℓ_2}, it is contained in P^{ℓ_2}, that is $y^* \leq y^{P,\ell_2}$. The vector of multipliers satisfies $\lambda \geq 0$ and thus $y^* = y^{P,\ell_2}$.

Definition 1 (*Heuristic CP method*) The constraint (50) imposes a search of paths focused on P^{ℓ} or alternatively in the graph associated with the nodes $\mathcal{N} - \mathcal{N}^{\ell}$. From a computational point of view it may be advantageous to force the exploration of P^{ℓ}. In this case the constraint (50) can be replaced by

$$\sum_{b \in \widehat{\mathcal{B}}^s} y_b \geq 1; \qquad s = 1, \ldots, \ell \tag{52}$$

and the resulting CP method is called a *heuristic CP* method because the convergence property given in Theorem 3 may be lost.

4 Computational experiments

In this section a set of numerical experiments was carried out with the aim of determining the suitability of the proposed methodology and analysing the obtained solutions. This section is organized into four subsections which deal with the following areas:

(i) Description of the test problem (in Sect. 4.1).
(ii) A set of numerical experiments to evaluate the use of CPLEX for solving [P1], [P2] and [P3] models (in Sect. 4.2).
(iii) A set of numerical tests to evaluate the Lagrangian relaxation for solving the [P3] model (in Sect. 4.3).
(iv) A sensitivity analysis for the parameter B_0 is presented in Sect. 4.4.

4.1 Description of the test problem

To test our proposed method we have used a real test problem from Castilla-La Mancha (Spain). The region of Castilla-La Mancha has an area of 79,226 km^2. We have considered various examples with 6, 67 and 290 towns with the highest populations and worked with the demands between these towns. Figure 4 shows the geographical region analysed. In this case the existing network $\mathcal{G}_A$ consists of 1,008 links which can be traversed in either direction. These links are classified as *highways* (average speed 112.5 km/h), *main roads* (average speed 100 km/h) and *local roads* (average speed 90 km/h). The network that can be built $\mathcal{G}_{\widehat{B}}$ comprises all the links which connect population centres not more than 30 km apart. This network contains 2,154 possible links.

Choosing this threshold guarantees that all highways are considered whose maximum distance between exits (associated to population centres) is less than 30 kms. Increasing this value would increase the cardinality of the set $\widehat{\mathcal{B}}$ and thus would increase the computational cost. The potential links are all the possible direct connections between population centres,

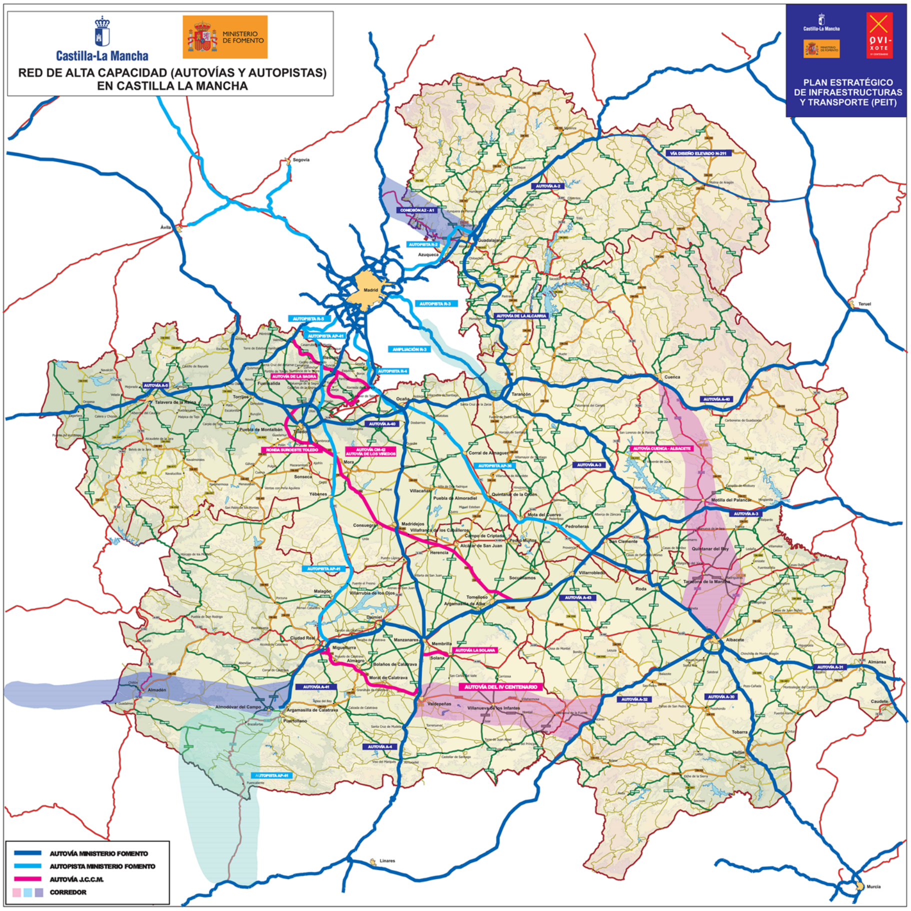


Fig. 4 Existing road network in Castilla-La Mancha (Spain)

whether or not they are currently covered by conventional roads. This means that an existing conventional road (joining two population centres directly) can be suitable for conversion to a highway, or if there is none it may be best to build a highway from scratch.

Two origin-destination matrices have been considered:

(g^1): This matrix has been calculated by the National Statistics Institute (INE) of Spain. It includes the daily journeys between population centres of Castilla-La Mancha.

(g^2): This O–D matrix was obtained from the synthetic distribution model with exponential decay function set out in Ortúzar and Willumsen (1994). The origin-destination matrix is calculated by:

$$g_{ij} \propto P_i P_j exp(-\beta c_{ij}),$$

where P_i and P_j are the populations of the centroids i and j, and c_{ij} is the distance between the two towns.

Three groups of centroids have been considered for the region of Castilla-La Mancha: CLM6 contains the 6 towns with more than 50,000 inhabitants, CLM67 contains the 67

Table 5 Size of the problem cases

Problem	Cities	$\mathcal{N}$	$\mathcal{A}$	$\widehat{\mathcal{B}}$	O–D Pairs (Case g^1)	O–D Pairs (Case g^2)
CLM6	6	290	1,008	2,154	36	36
CLM67	67	290	1,008	2,154	1,626	4,489
CLM290	290	290	1,008	2,154	7,244	84,100

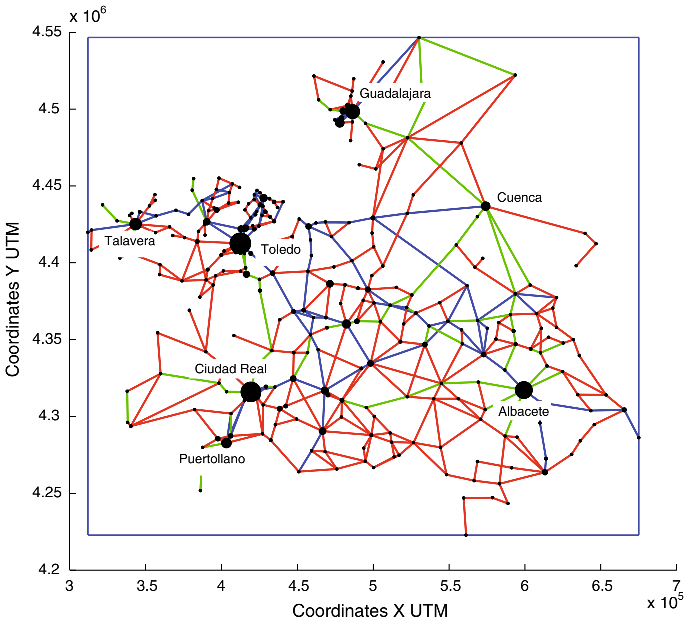


Fig. 5 Graph representing the existing road network $\mathcal{G}_A = (\mathcal{N}, \mathcal{A})$ (CLM6 test). (Color figure online)

towns with more than 5,000 inhabitants and CLM290 the 290 towns and villages with more than 1,000 inhabitants. Table 5 shows the size of the three problems.

Figures 5 and 6 show the CLM6 test. Figure 5 shows the existing road network $\mathcal{G}_A = (\mathcal{N}, \mathcal{A})$. The blue links represent the highways, the red links the main roads and the green links the local roads. This example only considers 6 towns (their names are shown) which can generate/attract demand, and the other towns and villages are considered for routing the demand appropriately. Figure 6 shows the network that could potentially be built $\mathcal{G}_B = (\mathcal{N}, \widehat{\mathcal{B}})$. The difference between the three tests is the set of O–D pairs considered.

The models have been solved with branch-and-cut algorithms performed in CPLEX 11 and integrated in GAMS 22.8. All the trials have been carried out with a Pentium computer with a 2.66 GHz Intel (R) Core(TM) 2 Quad CPU Q9450 microprocessor and 4 GB of RAM.

4.2 Solving [P1], [P2] and [P3] models with CPLEX

In this subsection, we show that, if the problem size is getting larger, it becomes impractical to solve the [P1], [P2] and [P3] models applying CPLEX in its original formulation with respect to storage requirement and computing time.

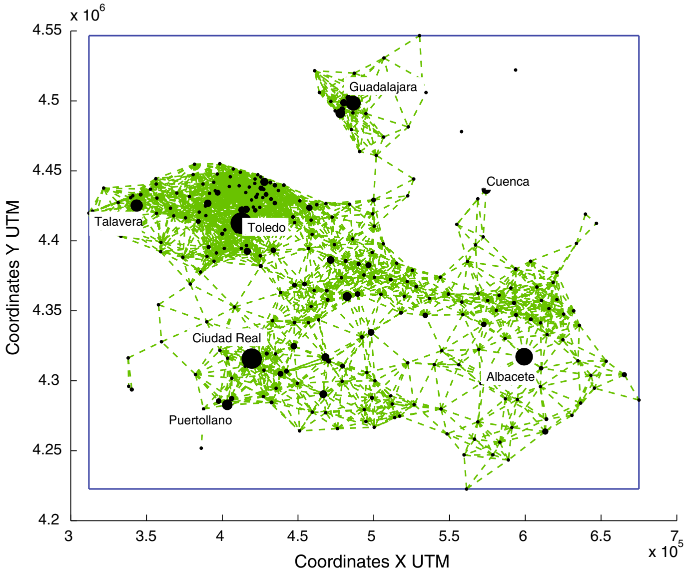


Fig. 6 Road network which may be built $\mathcal{G}_B = (\mathcal{N}, \widehat{\mathcal{B}})$

Table 6 Size statistics for [P1] model

	[P1]		[P1] based on O–D pairs (Case g^1)		[P1] based on O–D pairs (Case g^2)	
	Equations	Variables	Equations	Variables	Equations	Variables
CLM6	3,896	34,051	12,596	193,531	12,596	193,531
CLM67	21,586	358,327	111,098	8,645,971	1,303,966	23,865,679
CLM290	86,256	1,543,795	2,102,916	38,511,259	24,391,156	447,077,755

Furthermore, in relation to the storage requirement, we analyse the size of the [P1], [P2], [P2′] models in case studies CLM6, CLM67 and CLM290.

We notice that in the proposed formulation [P1] a commodity has multiple origins, so it can be treated as several commodities, each with a single origin and a single destination (i.e O–D pair commodities). It is a well known fact that a disaggregated formulation based on O–D pairs often gives a better performance, but its size may be huge. Table 6 compares the size of the [P1] model with respect to the disaggregated counterpart based on O–D pairs. The results show clearly that the number of variables and equations is huge, and it makes it impossible to address the disaggregated problems computationally.

Table 7 shows the size of the problems for the [P2] and [P2′] models. Both models have a similar size, although the [P2′] model has fewer equations than [P2].

The above tables show the order of magnitude of the problems. Next, we assess the computational burden. We have solved the [P1], [P2], [P2′] models with CPLEX. We have chosen the default settings in CPLEX and as the stopping criterion a relative gap lesser than 4 %.

Table 8 shows the CPU time (in seconds) and the relative gap for each [P1] problem test. The results show that the [P1] model is computationally tractable from a strategic point of

Table 7 Size statistics for [P2] and [P2′] models

	[P2]		[P2′]	
	Equations	Variables	Equations	Variables
CLM6	6,631	34,631	4,767	34,631
CLM67	24,321	358,907	22,457	358,907
CLM290	88,991	1,544,375	87,127	1,544,375

Table 8 CPU time (in seconds) for solving the [P1] models

	CPU time (Case g^1)	CPU time (Case g^2)
CLM6	11	2
	RGAP = 3.37 %	RGAP = 2.81 %
CLM67	431	6,679
	RGAP = 3.98 %	RGAP = 3.79 %
CLM290	184,639	282,120
	RGAP = 3.72 %	RGAP = 3.97 %

Table 9 CPU time (in seconds) for solving the [P2] and [P2′]models

	[P2]		[P2′]	
	CPU time (Case g^1)	CPU time (Case g^1)	CPU time (Case g^1)	CPU time (Case g^1)
CLM6	17	50	34	42
	RGAP = 3.33 %	RGAP = 3.39 %	RGAP = 3.95 %	RGAP = 3.91 %
CLM67	8,031	127,484	76,320	177,484
	RGAP = 2.85 %	RGAP = 7.11 %	RGAP = 4.38 %	RGAP = 7.13 %
CLM290	>20 days	>20 days	–	–
	RGAP = 4.79 %	RGAP = 8.11 %		

view, since the largest problem CLM290 takes around 51 and 78 h for the g^1 and g^2 matrices respectively. Table 9 depicts the same information as Table 8 for the [P2] and [P2′] models. The results show that the computational burden for the [P2] and [P2′] problems is significantly greater than for the [P1] problems but also the number of constraints and variables which are added to [P1] for obtaining the [P2] and [P2′] models are relatively small. Note that CPLEX was interrupted after 20 days for the problem CLM290 and the [P2] model. The solution obtained presented a relative gap of 4.79 and 8.11 % for g^1 and g^2 respectively. This experiment was not repeated for the [P2′] model because the problems would not be solved in a reasonable CPU time.

Moreover, the [P3] model introduces a set of constraints to avoid the sub-tours. The number of sub-tours in a graph $\mathcal{G}_B$ grows exponentially with the number of links. For these problems the number of links is $|\widehat{\mathcal{B}}| = 2{,}154$ and this fact and the above results show that the [P3] model for the real case study is too large for CPLEX.

4.3 Lagrangian relaxation for solving the [P3] model

In this subsection, the Lagrangian relaxation approach is assessed. The computational experiments were guided by three objectives: (1) to show that the case studies are soluble using

Table 10 CPU time (in seconds) for the [P3] models using the Lagrangian relaxation approach

	CPU time (Case g^1)		CPU time (Case g^1)	
	Approx. SUB_y	Accura. SUB_y	Approx. SUB_y	Accura. SUB_y
Heuristic CP method				
CLM6	11,860	92,470	24,730	122,331
	RGAP =4.83 %	RGAP =2.19 %	RGAP =0.59 %	RGAP =0.58 %
CLM67	17,430	160,900	25,690	142,890
	RGAP =3.92 %	RGAP =4.64 %	RGAP =2.49 %	RGAP =2.59 %
CLM290	21,100	–	27,950	–
	RGAP =1.91 %	RGAP = –%	RGAP =3.25 %	RGAP = –%
CP method				
CLM6	10,190	99,610	26,040	194,051
	RGAP =4.67 %	RGAP =3.12 %	RGAP =0.59 %	RGAP =0.83 %
CLM67	18,840	145,100	30,020	143,753
	RGAP =3.56 %	RGAP =3.24 %	RGAP =2.5 %	RGAP =3.24 %
CLM290	23,890	–	28,520	–
	RGAP =3.56 %	RGAP = –%	RGAP=2.89 %	RGAP= –%

the proposed Lagrangian relaxation, (2) to compare the *Heuristic* CP method with the CP method, by analysing the performance of each, and (3) to evaluate the accuracy in the solution of the sub-problems $\text{SUB}_y(\lambda)$ in the overall performance.

In order to achieve these objectives, we have run our problem tests with the *Heuristic* CP method and the CP method, using the following two strategies to solve $\text{SUB}_y(\lambda)$:

- Accura. SUB_y solves the sub-problems $\text{SUB}_y(\lambda)$ accurately using the parameter values $\varepsilon = 0.01$ and $\ell_{max} = 50$.
- Approx. SUB_y solves the sub-problems $\text{SUB}_y(\lambda)$ approximately using the parameter values $\varepsilon = 0.1$ and $\ell_{max} = 15$.

The performance of sub-gradient-based relaxation methods depends on the setting of several parameters. In this experiment we have omitted a fine-tuning for each problem through the analysis of a large number of strategies and parameter sets, and we have used general rules for setting their values. In all experiments we have set $\theta_\ell = 2$ for each iteration ℓ and the multipliers were initialized to $\lambda_b^0 = \alpha c_b$ with $\alpha = 0.16$. This choice is based on the interpretation of the multiplier λ_b as the travel time saved by a user when link b is built. Moreover, the sub-problems $\text{SUB}_x(\lambda)$ and $\text{SUB}_y(\lambda)$ were solved with CPLEX using the default settings. We have carried out a fixed number of iterations $\ell_{max}^{\text{outer}} = 30$ for all problem instances.

Table 10 shows the performance measures: (1) the relative GAP and (2) the CPU time. The conclusions drawn from the results in relation to the main objectives are:

- Both Heuristic CP and CP methods have a similar performance.
- A rough resolution strategy for the sub-problems $\text{SUB}_y(\lambda)$ is shown to be superior to an accurate resolution. It is important to mention that this strategy allows us to solve the case study within a time interval of 9 h.
- The computational burden is strongly influenced by the size of the design variable y and has a lesser effect on the number of commodities given, i.e on the flow variable x. Table 11 shows CPU time (in seconds) to solve sub-problems SUB_x and SUB_y.

Table 11 CPU time (in seconds) for solving sub-problems SUB_x and SUB_y

	CPU time (Case g^1)		CPU time (Case g^1)	
	`Approx.` SUB_y	`Accura.` SUB_y	`Approx.` SUB_y	`Accura.` SUB_y
Heuristic CP method				
CLM6	SUB_x 17.1	SUB_x 15.6	SUB_x 11.7	SUB_x 11.8
	SUB_y 11,840	SUB_y 92,456	SUB_y 24,720	SUB_y 122,320
CLM67	SUB_x 216	SUB_x 202	SUB_x 188	SUB_x 184
	SUB_y 17,211	SUB_y 160,730	SUB_y 25,500	SUB_y 142,706
CLM290	SUB_x 1,924	SUB_x –	SUB_x 1,661	SUB_x –
	SUB_y 19,808	SUB_y –	SUB_y 25,289	SUB_y –
CP method				
CLM6	SUB_x 15.9	SUB_x 16.7	SUB_x 11.9	SUB_x 14.6
	SUB_y 10,172	SUB_y 99,591	SUB_y 26,026	SUB_y 194,036
CLM67	SUB_x 198	SUB_x 231	SUB_x 220	SUB_x 234
	SUB_y 18,637	SUB_y 144,859	SUB_y 29,800	SUB_y 143,518
CLM290	SUB_x 1,921	SUB_x –	SUB_x 1,559	SUB_x –
	SUB_y 22,968	SUB_y –	SUB_y 26,960	SUB_y –

Figure 7 depicts the speed of convergence for the `CLM67` problem, illustrating the previous two conclusions.

4.4 Sensitivity analysis of parameter B_0

The parameter B_0 is the construction budget for the highway. This parameter determines the size of the feasible region and thus it is of great importance to the computational cost of any solution method. It is this question which is addressed in this section. Figure 10 shows the topology of the synthetic network (in four case studies) in which $\mathcal{G}_A$ has the layout of a grid. The network that can be built $\widehat{\mathcal{B}}$ comprises the diagonals of the squares. We consider the demand $g_\omega = 1$ between all the different pairs of nodes in the network, a link cost $c_a = 1$ for all $a \in \mathcal{A} \cup \mathcal{B}$ and a construction cost $f_b = 1$ for all $b \in \widehat{\mathcal{B}}$. The size of this problem consists of $|\widehat{\mathcal{B}}| = 162$ edges, $|\mathcal{N}| = 100$ nodes, $|\mathcal{A}| = 360$ links and $|\mathcal{W}| = 9{,}900$ O–D demand pairs.

We have solved [P3] problem by varying the values of $B_0 = 1, 2, \ldots, 10$ and performing $\ell_{max}^{\text{outer}} = 100$ main iterations of the CP method `Approx.` SUB_y with the same parameters as in the previous subsection. The obtained results are depicted in Fig. 8 which shows the relative GAP (%) and the objective function value (Z) versus the parameter B_0. Figure 9 shows CPU time (in seconds) versus B_0. It can be seen that an increase in the budget broadens the feasible region and so improves the solution (a reduction in total travel time in the network). We also see that the computational cost grows. This because when increasing the budget excessively, one part is used to locate the main highway, and the other part to locate cycles to serve other demands. This forces the CP algorithm to perform more iterations to eliminate all these cycles. The obtained solutions are shown in Fig. 10. The solutions for $B_0 = 1, 4, 7$ are on one of the diagonals and from arguments based on the geometry of the problem it is possible that these are optimum solutions.

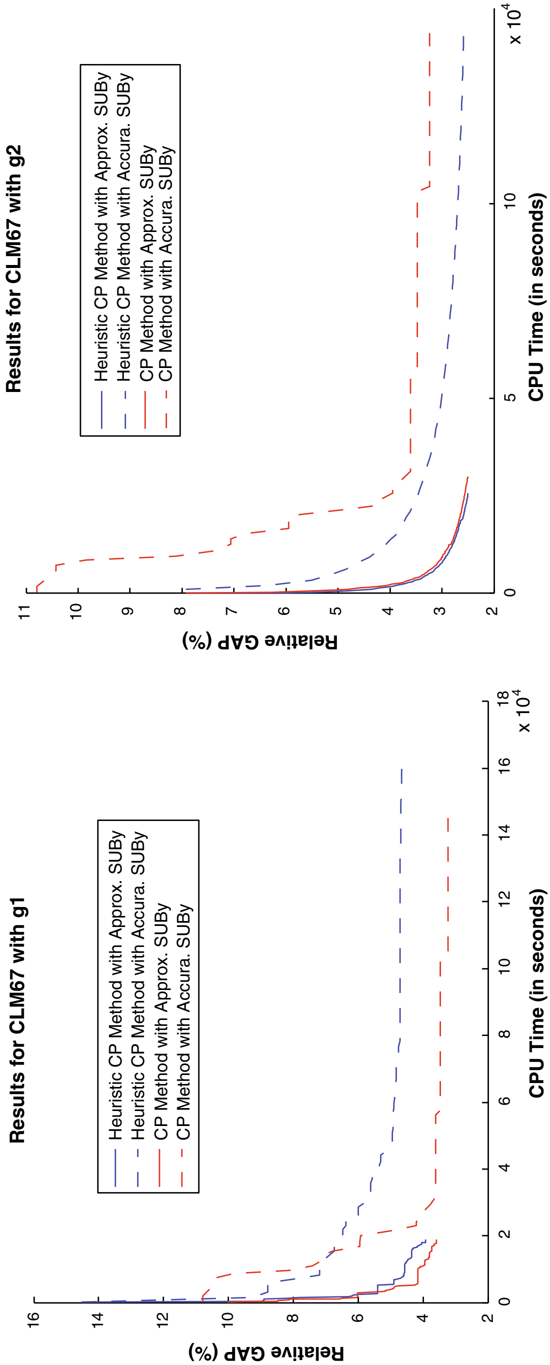


Fig. 7 Relative gap and CPU time (in seconds) for CLM67

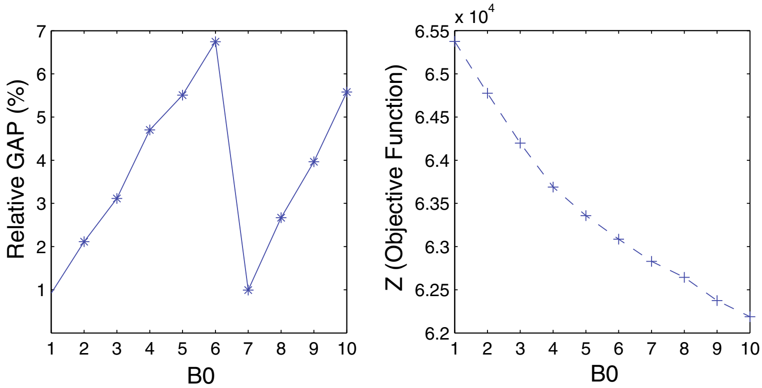
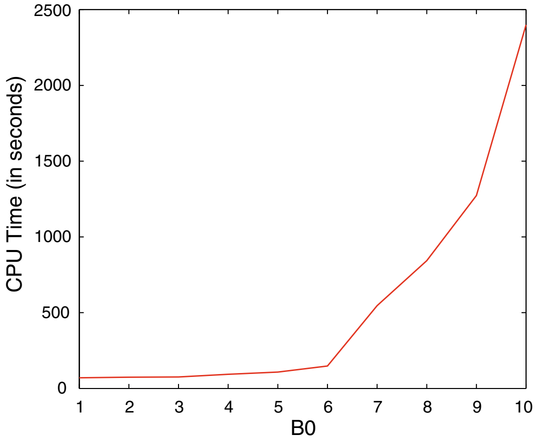


Fig. 8 Relative GAP and objective function versus B_0

Fig. 9 CPU time (in seconds) as a function of B_0



This example makes clear that in some applications it could be necessary to use bundle methods for solving the Lagrangian relaxation dual problem in order to reduce the number of main iterations.

5 Conclusions

We have analysed three MIP models, the so-called [P1], [P2] and [P3], for improving an existing road network. The [P3] model locates a highway defined as a path. This problem appears with other applications as the rapid transit network design, being computationally intractable using exact methods for large scale networks.

In this paper we propose a sub-gradient-based relaxation method for solving the [P3] model. The computational results suggest that the proposed method provides a means of obtaining near-optimal solutions for a real large scale application in the Castilla-La Mancha region (Spain).

The computational bottleneck of the sub-gradient method is the evaluation of the dual function ϕ. For this reason the success of an approach based on Lagrangian relaxation depends

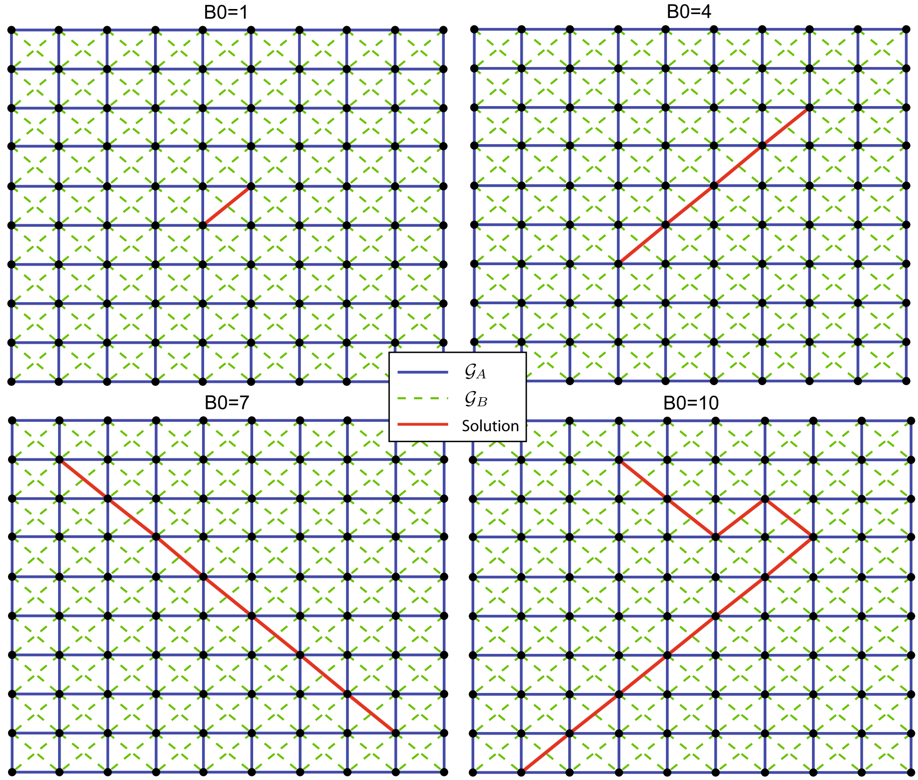


Fig. 10 Obtained solutions for different values of B_0

on the effective calculation of ϕ. Specifically, in this case, it relies on the effectiveness of the developed cutting-plane method. This fact justifies the application of a basic Lagrangian scheme without a fine-tuning of the parameters to the case study presented in this paper, testing whether an acceptable solution may be found. Once its effectiveness is demonstrated, bundle methods appear superior to sub-gradient approaches and can be used to speed up the proposed procedure.

In Reference Angulo et al. (2014) the problem of locating new highway corridors was also discussed. This approach requires 110 h to locate 3 highways in CLM67, which simultaneously takes into account microscopic cost structure (different in each geographical location). The approach described in this article has the advantage of considering in an unsimplified way the routing decisions of users. For the same problem the computational cost is 40 h but only one highway is located.

Acknowledgments This research has been financed by the *Ministerio de Economia y Competitividad* of Spain with the TRA-2011-27791-C03-03 research project.

References

Angulo, E., Castillo, E., García-Ródenas, R., & Sánchez-Vizcaíno, J. (2012). Determining highway corridors. *Journal of Transportation Engineering*, *138*(5), 557–570.

Angulo, E., Castillo, E., García-Ródenas, R., & Sánchez-Vizcaíno, J. (2014). A continuous bi-level model for the expansion of highway networks. *Computers & Operations Research*, *41*(2014), 262–276.
Boffey, B., & Narula, S. C. (1998). Models for multi-path covering-routing problems. *Annals of Operations Research*, *82*, 331–342.
Bruno, G., Gendreau, M., & Laporte, G. (2002). A heuristic for the location of a rapid transit line. *Computers & Operations Research*, *29*(1), 1–12.
Bruno, G., Ghiani, G., & Improta, G. (1998). A multi-modal approach to the location of a rapid transit line. *European Journal of Operational Research*, *104*(2), 321–332.
Contreras, I., Díaz, J. A., & Fernández, E. (2009). Lagrangean relaxation for the capacitated hub location problem with single assignment. *OR Spectrum*, *31*(3), 483–505.
Contreras, I., & Fernández, E. (2012). General network design: A unified view of combined location and network design problems. *European Journal of Operational Research*, *219*(3), 680–697.
Crainic, T. G., Frangioni, A., & Gendron, B. (2001). Bundle-based relaxation methods for multicommodity capacitated fixed charge network design. *Discrete Applied Mathematics*, *112*(1–3), 73–99.
Crainic, T. G., Gendreu, M., & Farvolden, J. M. (2000). A simplex-based tabu search method for capacitated network design. *INFORMS Journal on Computing*, *12*(3), 223–236.
Current, J., & Marsh, M. (1993). Multiobjective transportation network design and routing problems: Taxonomy and annotation. *European Journal of Operational Research*, *65*(1), 4–19.
Current, J., Pirkul, H., & Rolland, E. (1994). Efficient algorithms for solving the shortest covering path problem. *Transportation Science*, *28*(4), 317–327.
Current, J. R., ReVelle, C. S., & Cohon, J. L. (1984). The shortest covering path problem: An application of locational constraints to network design. *Journal of Regional Science*, *24*(2), 161–183.
Current, J. R., ReVelle, C. S., & Cohon, J. L. (1985). The maximum covering/shortest path problem: A multiobjective network design and routing formulation. *European Journal of Operational Research*, *21*(2), 189–199.
Current, J. R., ReVelle, C. S., & Cohon, J. L. (1986). The hierarchical network design problem. *European Journal of Operational Research*, *27*(1), 57–66.
Dufourd, H., Gendreau, M., & Laporte, G. (1996). Locating a transit line using tabu search. *Location Science*, *4*, 1–19.
Duin, C., & Volgenant, A. (1989). Reducing the hierarchical network design problem. *European Journal of Operational Research*, *39*(3), 332–344.
Erdemir, E. T., Batta, R., Spielman, S., Rogerson, P. A., Blatt, A., & Flanigan, M. (2008). Location coverage models with demand originating from nodes and paths: Application to cellular network design. *European Journal of Operational Research*, *190*(3), 610–632.
Escudero, L. F., & Muñoz, S. (2009). An approach for solving a modification of the extended rapid transit network design problem. *Top*, *17*(2), 320–334.
Fernández, P., & Marín, A. (2003). A heuristic procedure for path location with multisource demand. *INFOR: Information Systems and Operational Research*, *41*(2), 165–178.
Gutíerrez-Jarpa, G., Donoso, M., Obreque, C., & Marianov, V. (2010). Minimum cost path location for maximum traffic capture. *Computers & Industrial Engineering*, *58*, 332–341.
Holmberg, K., & Yuan, D. (2000). A lagrangian heuristic based branch-and-bound approach for the capacitated network design problem. *Operations Research*, *48*(3), 461–481.
Jha, M. K. (2001). Using a GIS for automated decision-making in highway cost analysis. *Transportation Research Record*, *1768*, 260–267.
Jha, M. K., McCall, C., & Schonfeld, P. (2001). Using GIS, genetic algorithms, and visualization in highway development. *Computer-Aided Civil and Infrastructure Engineering*, *16*(6), 399–414.
Jha, M. K., & Schonfeld, P. (2000). Integrating genetic algorithsms and geographic information system to optimize highway alignments. *Transportation Research Record*, *1719*, 233–240.
Jong, J. C. (1998). *Optimizing highway alignments with genetic algorithms*. Ph.D. thesis, College Park, University of Maryland
Kang, M. W. (2008). *An alignment optimization model for a simple highway network*. Ph.D. thesis, College Park, University of Maryland.
Konak, A. (2012). Network design problem with relays: A genetic algorithm with a path-based crossover and a set covering formulation. *European Journal of Operational Research*, *218*(3), 829–837.
Laporte, G., Marín, A., Mesa, J. A., & Ortega, F. A. (2007). *An integrated mehodology for the rapid transit network design problem. Lectures notes in computer science* (Vol. 4359, pp. 187–199).
Laporte, G., Marín, A., Mesa, J. A., & Perea, F. (2011). Designing robust rapid transit networks with alternative routes. *Journal of Advanced Transportation*, *45*(1), 54–65.
Laporte, G., Mesa, J. A., & Ortega, F. A. (2000). Optimization methods for the planning of rapid transit systems. *European Journal of Operations Research*, *122*(1), 1–10.

Laporte, G., Mesa, J. A., & Ortega, F. A. (2002). Locating stations on rapid transit lines. *Computers & Operations Research*, *29*(6), 741–759.

Laporte, G., Mesa, J. A., Ortega, F. A., & Perea, F. (2011). Planning rapid transit networks. *Socio-Economic Planning Sciences*, *45*(3), 95–104.

Laporte, G., Mesa, J. A., Ortega, F. A., & Sevillano, I. (2005). Maximizing trip coverage in the location of a single rapid transit alignment. *Annals of Operations Research*, *136*(1), 49–63.

Marín, A. (2007). An extension to rapid transit network design problem. *Top*, *15*, 231–241.

Marín, A., & García-Ródenas, R. (2009). Location of infrastructure in urban railway networks. *Computers & Operations Research*, *36*(5), 1461–1477.

Marín, A., Mesa, J. A., Perea, F. (2009). Integrating robust railway network design and line planning under failures. In R. K. Ahuja et al. (Ed.), *Robust and online large-scale optimization LNCS* (Vol. 5868, pp. 273–292).

Marín, A., & Pelegrín, B. (1999). Applying Lagrangian relaxation to the resolution of two-stage location problems. *Annals of Operations Research*, *86*, 179–198.

Mesa, J. A., & Boffey, T. B. (1996). A review of extensive facility location in networks. *European Journal of Operational Research*, *95*(3), 592–603.

Ortúzar, Jd D., & Willumsen, L. G. (1994). *Modelling transport*. Chichester: Wiley.

Pirkul, H., Current, J., & Nagarajan, V. (1991). The hierarchical network design problem: A new formulation and solution procedures. *Transportation Science*, *25*(3), 175–182.

Polyak, B. T. (1969). Minimization of unsmooth functionals. *USSR Computational Mathematics and Mathematical Physics*, *9*(3), 14–29.

Sancho, N. G. F. (1997). The hierarchical network design problem with multiple primary paths. *European Journal of Operational Research*, *96*(2), 323–328.

Solanki, R. S., Gorti, J. K., & Southworth, F. (1998). Using decomposition in large-scale highway network design with a quasi-optimization heuristic. *Transportation Research Part B: Methodological*, *32*(2), 127–140.

Timothy, J. (2013). *The maximal covering/shortest path problem revisited: An examination and reformulation of the problem to allow the elimination or attachment of sub-tours*. Ph.D. thesis, M. A. University of California, Santa Barbara

Yang, H., & Bell, M. G. H. (1998). Models and algorithms for road network design: A review and some new developments. *Transport Reviews: A Transnational Transdisciplinary Journal*, *18*(3), 257–278.