Communications in Combinatorics and Optimization
Vol. 11, No. 2 (2026), pp. 335-361
https://doi.org/10.22049/cco.2024.29105.1992

CCO
Commun. Comb. Optim.

Research Article

A traffic-based model to the p-median problem in congested networks

Mehdi Zaferanieh*, Maryam Abareshi†

Department of Applied Mathematics and Computer Science, Hakim Sabzevari University, Sabzevar, Iran
*m.zaferanieh@hsu.ac.ir
†abareshi66@gmail.com

Received: 6 February 2024; Accepted: 11 July 2024
Published Online: 18 July 2024

Abstract: In real urban transportation networks, the traffic counts of links affect their travel times, so considering fixed traffic-independent link travel times causes a lack of reliability in network models. In this paper, a traffic congestion model to the capacitated p-median problem is introduced to evaluate the effect of congestion on determining the optimal locations of facilities and allocating demands. Limited capacities for both nodes and links are considered, where increasing the traffic flow on one link would inevitably increase its travel time. Therefore, aside from considering the capacities of candidate points, the proposed model aims to locate facilities at nodes where their connected links also have enough capacities for passing through. Also, the effect of congestion imposed by the current flows corresponding to the existing origin-destination pairs in the network is considered. A generalized Benders decomposition algorithm is applied to reduce the problem to more manageable sub-problems, solved by the sub-gradient algorithm through consecutive iterations.

Keywords: network, capacitated p-median problem, traffic congestion, generalized Benders decomposition.

AMS Subject classification: 90B80

1. Introduction

The problem of locating p facilities in a plant or a network and allocating the demands of nodes to them is one of the basic ideas in location theory and management science. In the p-median problem the purpose is to locate p facilities so that the total sum of the

* *Corresponding Author*

weighted distances between demand nodes and their assigned facilities is minimized. Hakimi [22] provided the first study on the p-median problem, while the first linear integer mathematical model was presented by ReVelle and Swain [8, 27, 31]. Kariv and Hakimi [23] showed that the p-median problem on general graphs is NP-hard and presented an $O(p^2n^2)$ time algorithm to solve it on tree networks. Later, the time complexity of the algorithm on tree networks was improved to $O(pn^2)$ by Tamir [35]. For a complete bibliography on the capacitated and uncapacitated p-median problem where facilities are capacitated or uncapacitated, see [12, 15, 18, 24, 25, 30, 32]. The p-median problem would be more adapted to real applications if such affecting parameters as environmental and geographical conditions are considered in allocating demands, see [1, 38].

Assigning the demands of nodes to the existing facilities has a significant effect on traffic congestion, particularly when the capacities of links (roads) are limited [26]. Golabi et al. [21] investigated a discrete facility location problem with multiple servers in a congested network where the demands were uniformly distributed along the links. They intended to determine the number and the locations of facilities along with the number of assigned servers to each one in such a way that the traveling distance, the waiting time, the total cost, and the number of uncovered customers are minimized. Aboolian et al. [3] discussed a facility location problem in a congested network. The objective was to minimize the total fixed and service capacity costs, in such a way that the disutility related to the travel and waiting times for each node should be less than a predetermined threshold. Butun et al. [11] introduced a capacitated hub location and cargo routing problem considering the effect of congestion. Zaferanieh et al. [40] considered a *bi*-level p-median problem to choose the locations of some recreational facilities in $Mazandaran$ province, Iran. They considered additional criteria beyond distance that may influence customers' facility selection choices.

If we interpret the pairs of facilities and clients as origins and destinations, then some alternative mathematical models such as Traffic Assignment (TA), may be applied to the p-median problem to improve its reliability in terms of the travel times. In the TA model, first, some fixed values for travel demands between origin-destination ($O-D$) pairs are considered, then the route choices of users are determined with the best utility in terms of the travel cost. The TA problem is based on Wardrop's first principle [33, 36], which states "The journey times on all used routes are equal to or less than the journey time on any unused one" and results in the concept of user equilibrium (UE) on networks. The resulted network UE problem would be formulated as a mathematical model, including a convex nonlinear objective function as the sum of the integrals of the link performance functions, and a set of linear constraints [7].

Concerning the UE models, the users decide which route should be chosen. However, if the purpose is to minimize the total network cost (time), rather than considering the individual users' travel cost (time), the system-optimization (SO) model should be considered. In this case, Wardrop's second principle is considered, which states the average journey time of all users is minimal, [33, 36]. However, making all users sacrifice their interests to minimize the total network cost should be performed by an external factor. Pigou [29] noted the difference between the first and second Wardrops' principles and proposed differential taxation to divert traffic towards more efficient routes. Indeed, the difference between the UE and SO problems should be accounted for individual users to pay their contribution for the total travel cost, see [10, 13, 28].

In this paper, the effect of congestion on the p-median problem is examined by using an increasing link travel cost (time) function based on the traffic flow. It is supposed that a central unit is in charge of locating facilities, to minimize the total expenditure related to the establishment and transportation. Indeed, the model combines the p-median and the SO problems to locate p facilities and allocate the demands. However, to assure that the users in the network will behave according to the SO objective, it is necessary to add specified amounts as toll costs to the network links [10]. These values would be calculated by the obtained optimal solution of the SO problem, see [10]. The traffic-based p-median model introduced in this paper is more reliable compared to the classical p-median problem, particularly in the traffic-congested networks where the more link traffic counts, the higher would be their corresponding travel cost [26, 34, 39].

As an explicit application of the problem, consider a city with a predetermined set of $O - D$ pairs having specified flows to be traversed through their connecting paths. Under such circumstances, assume that the concerned authorities intend to establish p facilities, for example some storehouses or medical emergency centers, to serve the clients' demands in the network. As transportation from the demand nodes to the facilities (or vice versa) would inevitably impose additional traffic burden on the network links, we need a model to take it into consideration. Indeed, in addition to allocating demands to facilities and routing them, the preexisted flows in the network should be rerouted. The proposed model treats the problem in a SO approach to minimize the total establishment and transportation costs.

The mathematical model is a mixed-integer nonlinear programming ($MINLP$) problem for which a generalized Benders decomposition (GBD) method is applied. Using the GBD, the introduced traffic-based p-median problem is reduced to two linear binary and convex nonlinear sub-problems that would be solved, either by the branch and bound, or the sub-gradient algorithms. The lower and upper bounds are updated iteratively, and the convergence is verified in a finite number of iterations. The decomposition procedure was first developed by Benders [9] to solve mixed-variable programming problems. However, some restrictions regarding the convexity and other properties of the involved functions were considered [16]. The main contributions of this paper are listed as follows:

1. A traffic-based model for the p-median problem in congested networks is introduced, where the effect of congestion imposed by the flows of origin-destination pairs, and clients and facilities are simultaneously considered.

2. A tolling methodology is applied to adapt the UE flow pattern to the SO objective.

3. Using the GBD algorithm, the obtained $MINLP$ problem is reduced to a linear binary, and a nonlinear convex sub-problems in each iteration. The obtained sub-problems are solved by such appropriate methods as the branch and bound, and linearization approaches.

4. The convergence of the implemented algorithm is verified whereby the upper and lower bounds would converge together after a finite number of iterations.

5. Solving some numerical examples, and analyzing the obtained upper bounds and their corresponding solutions reveal the added value of the proposed nonlinear model and its solution approach.

The remainder of this paper is organized as follows. In Section (2), the proposed traffic-based p-median problem is presented. In Section (3), the general form of the GBD algorithm is introduced, while its application to the proposed problem is given in Section (4). Using the GBD, the problem is decomposed into more manageable sub-problems, and the upper and lower bounds are provided iteratively. The convergence of the GBD algorithm on the proposed problem is also demonstrated in Section (4). The efficiency of the model and the solution approach is investigated by some numerical examples in Section (5). Finally, the summary and conclusions are given in the last section.

2. Problem Formulation

Consider a network $G = (N, E)$ where $N = \{v_1, ..., v_n\}$ is the set of nodes and E is the set of links. Let $I \subseteq N$ denote the index set of candidate points for establishing facilities and $J \subseteq N$ represent the index set of demand (client) nodes. The purpose of the classic p-median problem is to select p nodes from the candidate set I and allocate the demand nodes to them in such away the total establishment and transportation cost is minimized. The frequently used notations are summarized in Table (1). Hereafter, for simplicity, we denote nodes v_i and v_j only by indices i and j. The classic capacitated p-median problem (CPM) is written as follows:

$$\min z(y, x) = \sum_{i\in I} e_i y_i + \sum_{i\in I}\sum_{j\in J} d_{ij} x_{ij} \tag{2.1}$$

$$s.t. \sum_{i\in I} y_i = p \tag{2.2}$$

$$\sum_{i\in I} x_{ij} = w_j, \ \forall j \in J \tag{2.3}$$

$$\sum_{j\in J} x_{ij} \leq a_i y_i, \ \forall i \in I \tag{2.4}$$

$$x_{ij} \geq 0, \quad y_i \in \{0, 1\}, \quad \forall \quad i \in I, j \in J.$$

Table 1. **The notation used in the p-median model**

Notation	Definition
N	$= \{v_1, \ldots, v_n\}$ The set of nodes
E	The set of links
I	The index set of candidate points for establishing facilities
J	The index set of demand (client) nodes
w_j	The demand of node $v_j,\ j \in J$
a_i	The specified capacity of node $v_i,\ i \in I$
e_i	The fixed set-up cost for establishing a facility at node $v_i,\ i \in I$
d_{ij}	The unit fixed cost of transporting demands from node $v_i,\ i \in I$ to node $v_j,\ j \in J$
Variables	
y_i	The binary decision variable indicating whether node v_i is selected to establish a facility or not
x_{ij}	The decision variable representing the amount of demand of node v_j provided by the facility at node v_i

The objective function (2.1) minimizes the cost of establishing facilities and serving client demands. Constraint (2.2) guarantees the number of established facilities is equal to p, while constraint (2.3) guarantees that the demands of all nodes $j,\ j \in J$ are met. Constraint (2.4) makes all open facilities serve the demands less than or equal to their capacities.

In the presence of congestion, the travel cost (time) on urban streets and intersections is an increasing function of their corresponding flow [33]. Consequently, a performance function, rather than a constant travel cost, should be associated with each link. As a result, the cost of transporting demands between nodes i and j depends on the traffic counts of links included in the traveled path. The link travel cost at zero flow is known as the free-flow travel cost, which is associated with the length of link [33]. As the flow increases, the travel cost also monotonically increases. The usual performance function for the travel cost (time) of links is the standard Bureau of Public Road (BPR) function defined by the following relation [2, 33]:

$$\tau_e(v_e) = t_e^0(1 + \beta(\frac{v_e}{c_e})^\alpha)$$

where the used parameters along with the additional notations required in the traffic-based model of the capacitated p-median problem are provided in Table (2). Also, α and β are fixed parameters which are usually equal to 4 and 0.15, respectively [2, 33]. Note that the values of the $O-D$ demands q_{rs} are usually available in a matrix called the $O-D$ trip matrix that could have been estimated by such alternative methods as the least squares or maximum entropy, see [2, 4]. Obviously, by adding the selected facilities to the network and routing their corresponding demands from different paths, the traffic counts on network links, and their travel cost will increase. As a result, the SO condition would be disrupted. Therefore, the previously existing $O-D$ flows should be rerouted to impose fewer travel costs, if possible. The traffic-based model of the capacitated p-median ($TCPM$) problem is introduced as follows:

Table 2. **The notation used in the traffic-based p-median model**

Notation	Definition
OD_{set}	The set of $O-D$ pairs
$\tau_e(.)$	The ascending travel cost function of link $e \in E$
t_e^0	The free-flow travel cost of link $e \in E$
c_e	The capacity of link $e \in E$
v_e	The traffic count of link $e \in E$
K_{ij}	The set of K-shortest paths connecting client-facility $(C-F)$ pair (i,j)
K'_{rs}	The set of K-shortest paths connecting $O-D$ pair (r,s)
q_{rs}	The given pre-specified $O-D$ demand of pair (r,s)
$\delta_{e,k}$	The binary parameter indicating whether the link e is a part of path k or not
Variables	
f_{ij}^k	The decision variable denoting the flow of path $k \in K_{ij}$
$g_{rs}^{k'}$	The decision variable representing the flow of path $k' \in K'_{rs}$

$$TCPM : \min z(y,f,g) = \sum_{i\in I} e_i y_i + \sum_{e\in E} v_e \tau_e(v_e) \tag{2.5}$$

$$s.t. \sum_{i\in I} y_i = p$$

$$\sum_{i\in I} x_{ij} = w_j, \ \forall j \in J$$

$$\sum_{j\in J} x_{ij} \le a_i y_i, \ \forall i \in I$$

$$\sum_{k\in K_{ij}} f_{ij}^k = x_{ij}, \quad \forall i \in I, \ j \in J \tag{2.6}$$

$$\sum_{k'\in K'_{rs}} g_{rs}^{k'} = q_{rs}, \quad \forall (r,s) \in ODset \tag{2.7}$$

$$\sum_{ij} \sum_{k\in K_{ij}} f_{ij}^k \delta_{e,k} + \sum_{rs} \sum_{k'\in K'_{rs}} g_{rs}^{k'} \delta_{e,k'} = v_e, \quad \forall e \in E \tag{2.8}$$

$$f_{ij}^k, g_{rs}^{k'} \ge 0, y_i \in \{0,1\}, \quad \forall \ \ i \in I, j \in J, k \in K_{ij}, (r,s) \in ODset, k' \in K'_{rs}.$$

In the $TCPM$, model (2.5), variables v_e are depended to both $g_{rs}^{k'}$ and f_{ij}^k, while variables x_{ij} are depended only to $g_{rs}^{k'}$. Therefore, the objective function would be served by variables y, f and g. The first part of the objective function (2.5) considers the cost of establishing facilities, while the second part includes the SO objective function regarding the total travel cost of links based on their traffic congestion. Constraint (2.6) imposes the amount of demand of node j which is served by the facility at node i equals the total sum of the demands traveling through different paths k from i to j. This constraint along with the nonnegativity constraints yields that $x_{ij} \ge 0$. Constraint (2.7) is explained similarly. Constraint (2.8) determines the traffic count of link e as the total sum of the flows of all paths passing through this link.

2.1. A brief discussion on congestion toll pricing

As mentioned before, to make network users behave jointly with the aim of minimizing the total network transportation cost, we need to set additional toll costs for traversing network links in such a way that the resulted UE solution is equivalent to the SO solution. Bergendorff et al. [10] showed that once the solution of the SO problem and the corresponding traffic counts of links are determined, the toll prices $\gamma_e,\ e \in E$ are calculated. Let v_e^{opt}s be the optimal traffic counts of links obtained by solving the SO model. Based on the method of Bergendorff et al. [10], if $\tau_e(.)$ is strictly monotonic and the amounts of $\nabla\tau_e(v_e^{opt})v_e^{opt}$ are nonnegative, the toll cost of link e is determined by $\gamma_e = \nabla\tau_e(v_e^{opt})v_e^{opt}$. Considering the definition of the BPR function, the cost map $\tau_e(.)$ clearly satisfies the desired conditions. Therefore, the imposed toll pricing method by Bergendorff et al. [10] is still valid for Problem (2.5).
The user equilibrium flow pattern with the link travel costs $\bar{\tau}_e(.) = \tau_e(.) + \gamma_e$, selected by the network users, matches that of the untolled SO problem, $\min_v \sum_{e\in E} v_e\tau_e(v_e)$. This is equivalent to the UE problem objective function $\min_v \sum_e \int_0^{v_e}(\tau_e(u) + \gamma_e)du$, see [10, 28]. Note that the cost of establishing facilities $\sum_i e_i y_i$ in the first part of the objective function is ineffective in this case. Indeed, the selected facilities act as some new destinations where their corresponding origins are the demand nodes. As a result, we solve the $TCPM$ problem (2.5) to find the optimal link traffic counts v_e^{opt}, and set a surplus cost $\gamma_e = \tau_e(v_e^{opt})v_e^{opt}$ for traversing along link e, to guarantee users will behave so that the SO objective function is minimized.

A Small Example. To compare the efficiency of the introduced $TCPM$ problem (2.5) with the classic CPM model (2.1), consider the small 4-node network depicted in Figure (1) where $I = J = N$. The free-flow travel costs, t_e^0, and capacities, c_e, of links are given next to them. It is assumed there is a self-loop for each node with $t_e^0 = 0$ and $c_e = M$ where M is a large number. In addition, the demand of each node w_j, as well as its capacity a_i, when selected as a facility, are given in Figure (1). The set-up fixed costs for all nodes are considered equal to 20. Finally, the $O - D$ trip matrix related to the existing flows is $A = \begin{bmatrix} 0 & 15 & 8 & 7 \\ 10 & 0 & 14 & 20 \\ 15 & 25 & 0 & 12 \\ 20 & 10 & 10 & 0 \end{bmatrix}$.

We have solved the $TCPM$ model (2.5) and the classic CPM model (2.1) for $p = 1$ by using the existing mixed-integer solvers in $LINGO$ 18.0. By applying the fixed costs t_e^0, the costs of shortest paths from each node i to all other nodes j are calculated and used as the values of d_{ij}. Note that, due to one-way links, d_{ij} differs from d_{ji}. The 1-median solution obtained by the CPM model (2.1) is node 4 with the objective function value 670.

On the other hand, considering the flow-dependent link travel costs $\tau_e(.)$, the solution to the $TCPM$ model (2.5) is determined at node 1 with the objective value 6983.70.

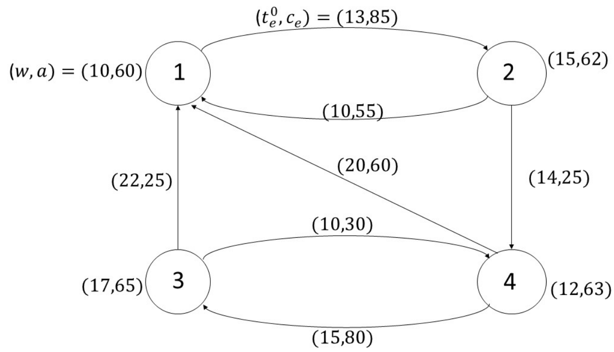

Figure 1. Small network

If capacities of all links are changed to $M = 1000$, and the effect of existing $O-D$ flows is ignored, then the 1-median solution of the $TCPM$ problem moves to node 4 with the objective function 670. Indeed, considering only the fixed free-flow travel costs of links, the total sum of weighted distances from node 4 to all other nodes is less than that of node 1, still the small capacities of links entering node 4 and the congestion around, prevent it from being selected as the facility in the $TCPM$ model. The CPM problem (2.1) is a relaxed version of the introduced $TCPM$ model (2.5), wherein the effects of link capacities and existing $O-D$ flows in the network are ignored. Indeed, the $TCPM$ model (2.5) benefits from greater reliability in real networks where the traffic congestion increases the transportation costs. In the following sections, first referring to [5], the general implementation of the GBD algorithm on $MINLP$ models is stated. Then, the GBD algorithm is applied to transform the proposed nonlinear mixed-integer $TCPM$ model (2.5) into two convex, and linear mixed-binary sub-problems by which the upper and lower bounds are updated iteratively, see [5].

3. Generalized Benders Decomposition Algorithm

The decomposition method was first developed by Benders [9] and extended by Geoffrion [20]. Geoffrion [20] suggested the GBD algorithm to solve specific nonlinear programming (NLP) and $MINLP$ problems. Then, Bagajewicz and Manousiouthakis [5] applied the GBD algorithm to solve the general form of $MINLP$ models. Recently, Zaferanieh [37] applied the GBD algorithm to solve a most probable facility location problem where there are some more criteria, aside form the transportation cost, for determining the allocation solution. To explain the outline of the GBD

method, consider the following model:

$$\min_{x,y} F(x,y) \tag{3.1}$$
$$s.t.\;\; G(x,y) \leq 0,$$
$$x \in X, \quad y \in Y.$$

Geoffrion [20] proposed to consider the projection of the above problem on Y as:

$$\min_{y} v(y) \tag{3.2}$$
$$s.t.\;\; v(y) = \min_{x\in X}\{F(x,y); \;\; s.t.\; G(x,y) \leq 0\}$$
$$y \in Y \cap V,$$

where $V = \{y : \; G(x,y) \leq 0 \text{ for some } x \in X\}$. Geoffrion [20] proposed a decomposition of (3.2) when X is a convex set, and the functions $F(x,y)$ and $G(x,y)$ are convex with respect to the variable x. Take the following two problems:

[1.] Primal problem:

$$\min_{x} F(x,\bar{y}), \tag{3.3}$$
$$s.t.\;\; G(x,\bar{y}) \leq 0$$
$$x \in X,$$

in which $\bar{y}$ is an arbitrary fixed point in Y.

[2.] Master problem:

$$\min_{y\in Y}\{\max_{u\geq 0}\{\min_{x\in X} F(x,y) + u^t G(x,y)\}\} \tag{3.4}$$
$$s.t.\;\; \min_{x\in X}\{\lambda^t G(x,y)\} \leq 0, \; \forall \lambda \in \Lambda,$$

where $\Lambda = \{\lambda \in R^m; \lambda \geq 0, \;\; \sum_{i=1}^m \lambda_i = 1\}$ (m is the size of vector G).

The master problem is equivalent to the projection problem (3.2), when X is a convex set, and $F(x,y)$ and $G(x,y)$ are convex functions with respect to the variable x, see [5]. In addition, problem (3.4) can be reformulated as follows:

$$\min_{y,y_0} y_0 \tag{3.5}$$
$$s.t.\;\; L^*(y,u) = \min_{x\in X}\{F(x,y) + u^t G(x,y)\} \leq y_0, \quad \forall u \geq 0 \tag{3.6}$$
$$L_*(y,\lambda) = \min_{x\in X}\{\lambda^t G(x,y)\} \leq 0, \quad \forall \lambda \in \Lambda. \tag{3.7}$$

Geoffrion [20] suggested solving a relaxed version of problem (3.5), in which all but a few constraints are ignored. Indeed, in each iteration, it is required to take the constraints corresponding only to a subset of dual variables u and λ, while the others are put aside. The obtained problem is called the relaxed master problem and solved iteratively to add the optimality and feasibility constraints [5]. Therefore, the optimal values of the objective function would provide a monotone nondecreasing sequence as the lower bounds to the problem (3.1). The steps of the *GBD* algorithm in the general form are represented in Algorithm (1).

Algorithm 1 The general form of the *GBD* algorithm

Step 1. Consider a feasible point $\bar{y} \in Y \cap V$. Solve the primal problem (3.3) to obtain the optimal solution x^* and the optimal multiplier vector u^*. Set $K^{feas} = 1$, $K^{infeas} = 0$, $u^{(1)} = u^*$ and $UB = F(x^*, \bar{y})$. Select a tolerance $\epsilon > 0$ and determine the function $L^*(y, u^{(1)})$.

Step 2. Obtain the global optimal solution of the current relaxed master problem:

$$\begin{aligned} &\min_{y, y_0} y_0 \\ &s.t.\ \ L^*(y, u^{(k_1)}) \le y_0,\ k_1 = 1, ..., K^{feas}, \quad \text{(optimality constraint)} \\ &\qquad\qquad L_*(y, \lambda^{(k_2)}) \le 0,\ \ k_2 = 1, ..., K^{infeas}, \quad \text{(feasibility constraint).} \end{aligned}$$

Let $(\hat{y}, \hat{y}_0)$ be the global optimal solution. Set $LB = \hat{y}_0$. If $UB \le LB + \epsilon$, terminate.

Step 3. Examine the feasibility of the primal problem (3.3) corresponding to $\bar{y} = \hat{y}$.

Step 3.*a*. Feasible primal problem: Solve the primal problem (3.3) by using $\bar{y}$. If $v(\bar{y}) \le LB + \epsilon$, terminate. Otherwise, determine the optimal multiplier vector u^*, set $K^{feas} = K^{feas} + 1$ and $u^{(K^{feas})} = u^*$. If $v(\bar{y}) < UB$, replace UB with $v(\bar{y})$. Finally, determine the function $L^*(y, u^{(K^{feas})})$ and return to Step 2.

Step 3.*b*. Infeasible primal problem: Determine a set of values $\lambda^* \in \Lambda$ satisfying $\min_{x \in X}\{(\lambda^*)^t G(x, \bar{y})\} > 0$ by solving problem (3.8):

$$\min_{x, \alpha} \alpha \tag{3.8}$$

$$s.t.\ \ G(x, \bar{y}) - \alpha \underline{1} \le 0, \tag{3.9}$$

$$x \in X,\ \alpha \in R,$$

where $\underline{1} = (1, 1, ..., 1)^t$. The optimal dual multiplier vector to constraint (3.9) provides the desired λ^*. Set $K^{infeas} = K^{infeas} + 1$ and $\lambda^{(K^{infeas})} = \lambda^*$. Determine the function $L_*(y, \lambda^{(K^{infeas})})$. Go back to Step 2.

Bagajewicz and Manousiouthakis [5] studied the importance of two properties, *Property (P)* and *L-Dual-Adequacy* defined by Geoffrion [20], for proper application of the algorithm. They showed that if *Property (P)* is not met, the algorithm may result in points that are not even locally optimal. In addition, violating the requirement of *L-Dual-Adequacy* may also lead to local optimal points or solutions that are not even extremal.

Briefly, *Property (P)* states that for every $u \geq 0$ and $\lambda \in \Lambda$, the global minimums in constraints (3.6) and (3.7) over X can be obtained independently of y. Therefore, given the dual multipliers u and λ, the functions $L^*(y, u)$ and $L_*(y, \lambda)$ are explicitly determined. Once *Property (P)* holds, the evaluation of $L^*(y, u)$ and $L_*(y, \lambda)$ requires the minimums in (3.6) and (3.7) to be global. This property is called *L-Dual-Adequacy*, and holds when a global search is conducted to obtain the solutions of $\min_{x \in X}\{F(x, y) + u^t G(x, y)\}$ and $\min_{x \in X}\{\lambda^t G(x, y)\}$, see [5, 20].

Remark 1. If functions $F(x, y)$ and $G(x, y)$ are separable in x and y and convex on set X, the global minimums over x in constraints (3.6) and (3.7) are obtained independently of y, see [5]. Therefore, functions $L^*(y, u^{(K^{feas})})$ and $L_*(y, \lambda^{(K^{infeas})})$ would be defined explicitly during the steps of Algorithm (1) .

If problem (3.1) possesses certain conditions, the *GBD* algorithm will converge in a finite number of iterations. Indeed, the smallest value of the primal objective function and the global solution of the master problem come close together; therefore, the global optimum of the overall problem would be provided within a pre-specified tolerance, see [20]. However, when convexity on x does not hold, dual gap may exist. Nevertheless, the global solution to the master problem will still provide a valid lower bound to problem (3.1). Next, the application of the *GBD* algorithm on the *TCPM* problem (2.5) is investigated, and the convergence conditions are verified.

4. Implementing the *GBD* algorithm on the *TCPM* problem

To apply the proposed *GBD* algorithm, we first reformulate the *TCPM* problem (2.5) as follows:

$$\min z(y, f, g) = \sum_{i \in I} e_i y_i + \sum_{e \in E} v_e \tau_e(v_e) \tag{4.1}$$

$$\sum_{j \in J} x_{ij} \leq a_i y_i, \ \forall i \in I \tag{4.2}$$

$$f, g \in S, \ y \in Y,$$

where $S = \{f_{ij}^k, g_{rs}^{k'} \geq 0, \ \sum_{i \in I} x_{ij} = w_j, \ \sum_{k \in K_{ij}} f_{ij}^k = x_{ij}, \ \sum_{k' \in K'_{rs}} g_{rs}^{k'} = q_{rs}, \ \sum_{ij} \sum_{k \in K_{ij}} f_{ij}^k \delta_{e,k} + \sum_{rs} \sum_{k' \in K'_{rs}} g_{rs}^{k'} \delta_{e,k'} = v_e, \forall i, j, k, (r, s), k', e\}$ and $Y = \{y; \ y_i$ is binary and $\sum_{i \in I} y_i = p\}$. Note that, by defining the performance function $\tau_e(.)$ based on the *BPR* function, both $z(y, f, g)$ and the $h_i(y, f, g) = \sum_{j \in J} x_{ij} - a_i y_i, \ \forall i \in I$ are convex for variables $f_{ij}^k, g_{rs}^{k'}$ on convex set S. Following the steps of Algorithm (1), the implementation of the *GBD* algorithm to solve the *TCPM* problem is summarized in Algorithm (2).

Note that in problem (4.6), the variables $u^{(k_1)}, x^{(k_1)}$, $\lambda^{(k_2)}$ and $\tilde{x}^{(k_2)}$ are specified fixed values that have been obtained during the previous steps of the algorithm;

Algorithm 2 The GBD algorithm on the $TCPM$ problem

Step 1 Choose a tolerance $\epsilon > 0$ and a feasible vector $\bar{y} \in Y$ for which there is a feasible solution vector $s = (f, g) \in S$ satisfying in (4.2).

Step 1.a Solve the corresponding primal problem:

$$\min_{f,g \in S} z(\bar{y}, f, g) = \sum_{i \in I} e_i \bar{y}_i + \sum_{e \in E} v_e \tau_e(v_e) \tag{4.3}$$

$$\sum_{j \in J} x_{ij} \leq a_i \bar{y}_i, \ \forall i \in I, \tag{4.4}$$

and obtain the global optimal solution $s^* = (f^*, g^*)$, and multipliers $u_i^*,\ i \in I$ corresponding to constraints (4.4). Set $K^{feas} = 1,\ K^{infeas} = 0,\ UB = z(\bar{y}, f^*, g^*)$ and vector $u^{(1)} = u^*$.

Step 1.b Solve the following sub-problem (4.5) to obtain the optimal solution $\hat{s} = (\hat{f}, \hat{g})$:

$$\min_{f,g \in S} \sum_{e \in E} v_e \tau_e(v_e) + \sum_{i \in I} u_i^{(K^{feas})} \sum_{j \in J} x_{ij}, \tag{4.5}$$

then, set $x_{ij}^{(K^{feas})} = \sum_{k \in K_{ij}} \hat{f}_{ij}^k,\ \forall i \in I,\ j \in J$ and $v_e^{(K^{feas})} = \sum_{ij} \sum_{k \in K_{ij}} \hat{f}_{ij}^k \delta_{e,k} + \sum_{rs} \sum_{k' \in K'_{rs}} \hat{g}_{rs}^{k'} \delta_{e,k'},\ \forall e \in E$.

Step 2 Solve globally the current relaxed master problem:

$$\min_{y, y_0 \in Y} y_0 \tag{4.6}$$

$$s.t.\ \ L^*(y, u^{(k_1)}) = \sum_{i \in I} e_i y_i + \sum_{e \in E} v_e^{(k_1)} \tau_e(v_e^{(k_1)}) + \sum_{i \in I} u_i^{(k_1)} (\sum_{j \in J} x_{ij}^{(k_1)} - a_i y_i) \leq y_0, \quad k_1 = 1, ..., K^{feas},$$

$$L_*(y, \lambda^{(k_2)}) = \sum_{i \in I} \lambda_i^{(k_2)} (\sum_{j \in J} \tilde{x}_{ij}^{(k_2)} - a_i y_i) \leq 0, \quad k_2 = 1, ..., K^{infeas},\ \lambda \in \Lambda.$$

The global optimal solution to the linear mixed-binary model (4.6) could be obtained by such efficient methods as the branch and bound. Let $(\hat{y}, \hat{y}_0)$ be the global optimal solution. Set $LB = \hat{y}_0$. If $UB \leq LB + \epsilon$, terminate.

Step 3. Examine the feasibility of the primal problem (4.3) corresponding to $\bar{y} = \hat{y}$ by solving the following problem:

$$\min_{f,g,\alpha} \alpha$$

$$\sum_{j \in J} x_{ij} - a_i \bar{y}_i - \alpha \leq 0, \quad \forall i \in I \tag{4.7}$$

$$f, g \in S,\ \ \alpha \in R.$$

Step 3.a If $\alpha \leq 0$, then the primal problem (4.3) is feasible. Find the optimal solution $s^* = (f^*, g^*)$ of model (4.3) corresponding to the given vector $\bar{y}$. If $z(\bar{y}, f^*, g^*) \leq LB + \epsilon$, terminate. Otherwise, determine the optimal multiplier vector u^* corresponding to constraints (4.4), set $K^{feas} = K^{feas} + 1$ and $u^{(K^{feas})} = u^*$. If $z(\bar{y}, f^*, g^*) < UB$, put $UB = z(\bar{y}, f^*, g^*)$. Finally, return to Step 1.b

Step 3b. If $\alpha > 0$, then the primal problem (4.3) is infeasible. Determine the optimal multiplier vector λ^* corresponding to constraint (4.7). Set $K^{infeas} = K^{infeas} + 1$ and $\lambda^{(K^{infeas})} = \lambda^*$. Find the optimal solution $\tilde{s} = (\tilde{f}, \tilde{g})$ to the following model:

$$\min_{f,g \in S} \sum_{i \in I} \lambda_i^{(K^{infeas})} \sum_{j \in J} x_{ij}. \tag{4.8}$$

Set $\tilde{x}_{ij}^{(K^{infeas})} = \sum_{k \in K_{ij}} \tilde{f}_{ij}^k, \ \forall i \in I, \ j \in J$ and go to Step 2.

therefore, the problem (4.6) is a linear mixed-binary model. Next, the convergence conditions of Algorithm (2), stated by Geoffrion [20], are verified for problem (2.5).

Theorem 1. *The generalized Benders decomposition Algorithm (2) for the TCPM problem (2.5) terminates in a finite number of iterations for any given $\epsilon > 0$.*

Proof. Since $Y = \{y; y_i$ is binary and $\sum_{i \in I} y_i = p\}$ is a finite discrete set, Algorithm (2) would be convergent if the following conditions hold, see [20]:

i. S is a nonempty convex set. Further, the set $B_y = \{B = (b_1, \ldots, b_{|I|}) : \ h_i(y, f, g) \leq b_i$ for some $s = (f, g) \in S\}$ is closed for each fixed $y \in Y$.

ii. Functions $z(y, f, g)$ and $h_i(y, f, g)$ are convex on S for each fixed $y \in Y$. In addition, for each fixed $\bar{y} \in Y \cap V$, the optimal value $z(\bar{y}, f, g)$ is infinite, or the primal problem (4.3) possesses an optimal multiplier vector.

Part *i*. The set S is nonempty and convex; because setting $f_{jj}^k = w_j$ for a specified path $k \in K_{jj}$ and $g_{rs}^{k'} = q_{rs}$ for a specified path $k' \in K'_{rs}$ together with setting zero flows for other paths, provides a feasible solution. In addition, S is closed and bounded. Since functions $h_i(y, f, g)$ are continuous on S, the set B_y is also closed, see [20].

Part *ii*. For each fixed y, the convexity of functions $z(y, f, g) = \sum_{i \in I} e_i y_i + \sum_{e \in E} v_e \tau_e(v_e)$ and $h_i(y, f, g) = \sum_{j \in J} x_{ij} - y_i a_i$ with respect to variables f, g is clear by the definition of $\tau_e(.)$. Following the saddle point optimality conditions stated in [6], the existence of the optimal multiplier vector is inferred, see [19]. □

In the following, we investigate the solution approach to the nonlinear problem (4.3) obtained from Step 1.a of Algorithm (2) for a given vector $\bar{y}$ of selected facilities. Indeed, model (4.3) is a convex nonlinear problem for which the convex combination method is proposed.

Convex combination method to the primal problem

By removing the constant values of the objective function, the primal problem (4.3) to be solved in Step 1.a of Algorithm (2) for a given vector $\bar{y}$ is written as follows:

$$\min z_{prim}(\bar{y}, f, g) = \sum_{e\in E} v_e \tau_e(v_e) \tag{4.9}$$

$$\sum_{i\in I} x_{ij} = w_j, \ \forall j \in J \tag{4.10}$$

$$\sum_{j\in J} x_{ij} \leq a_i \bar{y}_i, \ \forall i \in I \tag{4.11}$$

$$\sum_{k'\in K'_{rs}} g_{rs}^{k'} = q_{rs}, \quad \forall (r,s) \in ODset$$

$$\sum_{k\in K_{ij}} f_{ij}^{k} = x_{ij}, \quad \forall i \in I, \ j \in J$$

$$\sum_{ij} \sum_{k\in K_{ij}} f_{ij}^{k} \delta_{e,k} + \sum_{rs} \sum_{k'\in K'_{rs}} g_{rs}^{k'} \delta_{e,k'} = v_e, \ \forall e \in E$$

$$f_{ij}^{k}, g_{rs}^{k'} \geq 0, \forall \qquad i \in I, j \in J, k \in K_{ij}, (r,s) \in ODset, \ k' \in K'_{rs}.$$

The nonlinear convex problem (4.9) is similar to the SO model with a marginal difference [33]. Indeed, in the SO problem, the amounts of serviced $C-F$ demands x_{ij} are available previously and constraints (4.10), and (4.11) are omitted. We suggest using the convex combination method, initially proposed by Frank and Wolf [17] for solving quadratic programming problems with linear constraints. The so-called Frank-Wolf (FW) algorithm is a feasible descent direction method, starting with a given initial solution, where by applying the first-order approximation of the objective function, the descent direction is derived iteratively. Then, the step size along the descent direction is sought by minimizing a univariable problem. For a convex differentiable objective function and a compact feasible area, the rate of convergence of this method is $O(1/\epsilon)$, [2, 17, 33].

Turning back to problem (4.9), the linear program which has to be solved in the direction-finding step of the FW algorithm, written in terms of path flow variables $f_{ij}^{k}, g_{rs}^{k'}$, at the nth iteration becomes as follows:

$$\min z_{lin}(f, g) = \sum_{ij} \sum_{k\in K_{ij}} (\sum_{e} (\tau_e(v_e^n) + \tau'_e(v_e^n) v_e^n) \delta_{e,k}) f_{ij}^{k} \tag{4.12}$$

$$+ \sum_{rs} \sum_{k'\in K'_{ij}} (\sum_{e} (\tau_e(v_e^n) + \tau'_e(v_e^n) v_e^n) \delta_{e,k'}) g_{rs}^{k'} \tag{4.13}$$

$$\sum_{i\in I} \sum_{k\in K_{ij}} f_{ij}^{k} = w_j, \ \forall j \in J$$

$$\sum_{k'\in K'_{rs}} g_{rs}^{k'} = q_{rs}, \quad \forall (r,s) \in ODset$$

$$\sum_{j\in J}\sum_{k\in K_{ij}} f_{ij}^k \leq a_i\bar{y}_i,\ \forall i\in I$$
$$f_{ij}^k, g_{rs}^{k'} \geq 0, \forall \qquad i\in I, j\in J, k\in K_{ij}, (r,s)\in ODset,\ k'\in K'_{rs}.$$

In the above linear problem $v_e^n = \sum_{ij}\sum_{k\in K_{ij}} f_{ij}^{k,n}\delta_{e,k} + \sum_{rs}\sum_{k'\in K'_{rs}} g_{rs}^{k',n}\delta_{e,k'}$ where $f_{ij}^{k,n}, g_{rs}^{k',n}$ are the estimated path flow variables in the nth iteration. Note that the objective function consists of the derivatives of the function $z_{prim}(\bar{y}, f, g)$ with respect to variables f_{ij}^k and $g_{rs}^{k'}$ calculated as $\dfrac{\partial z_{prim}(\bar{y},f,g)}{\partial f_{ij}^k} = \sum_e \dfrac{\partial z_{prim}(\bar{y},f,g)}{\partial v_e}\delta_{e,k} = \sum_e(\tau_e(v_e)+\tau'_e(v_e)v_e)\delta_{e,k}$ and $\dfrac{\partial z_{prim}(\bar{y},f,g)}{\partial g_{rs}^{k'}} = \sum_e \dfrac{\partial z_{prim}(\bar{y},f,g)}{\partial v_e}\delta_{e,k'} = \sum_e(\tau_e(v_e)+\tau'_e(v_e)v_e)\delta_{e,k'}$. The steps of the convex combination method to solve the convex problem (4.9) is summarized in Algorithm (3).

Algorithm 3 Convex combination method

Step 0: Initialization. Solve problem (4.13) with $f_{ij}^{k,0} = 0,\ \forall i,j,k$ and $g_{rs}^{k',0} = 0,\ \forall (r,s), k'$, i.e. $v_e^0 = 0$ for all $e\in E$. Set counter $n = 1$ and name the obtained optimal solution as $f_{ij}^{k,n}, g_{rs}^{k',n}$.

Step 1: Calculate $v_e^n = \sum_{ij}\sum_{k\in K_{ij}} f_{ij}^{k,n}\delta_{e,k} + \sum_{rs}\sum_{k'\in K'_{rs}} g_{rs}^{k',n}\delta_{e,k'}$ and solve problem (4.13) to obtain the optimal solution $\hat{f}_{ij}^k, \hat{g}_{rs}^{k'}$.

Step 2: Set direction $\eta_{ij}^{k,n} = \hat{f}_{ij}^k - f_{ij}^{k,n},\ \rho_{rs}^{k',n} = \hat{g}_{rs}^{k'} - g_{rs}^{k',n}$. Solve the following line search problem to find the step size β:

$$\min z_{step}(\beta) = \sum_{e\in E} v_e^\beta \tau_e(v_e^\beta)$$
$$s.t.\quad 0\leq\beta\leq 1.$$

where $v_e^\beta = \sum_{ij}\sum_{k\in K_{ij}}(f_{ij}^{k,n} + \beta\eta_{ij}^{k,n})\delta_{e,k} + \sum_{rs}\sum_{k'\in K'_{rs}}(g_{rs}^{k',n} + \beta\rho_{rs}^{k',n})\delta_{e,k'}$.

Step 4: Move. Set $f_{ij}^{k,n+1} = f_{ij}^{k,n} + \beta\eta_{ij}^{k,n}$ and $g_{rs}^{k',n+1} = g_{rs}^{k',n} + \beta\rho_{rs}^{k',n}$ for all i,j,k and $(r,s), k'$.

Step 5: Convergence test. For a given $\epsilon > 0$, if $|z_{prim}(\bar{y}, f^n, g^n) - z_{prim}(\bar{y}, f^{n+1}, g^{n+1})| \leq \epsilon$, stop; the current solutions $f_{ij}^{k,n+1}, g_{rs}^{k',n+1}$ are the optimal path flows. Otherwise, set $n = n+1$ and go to Step 1.

In the next section, the added value of the proposed model and the solution algorithm is verified via some numerical examples.

5. Numerical examples

In this section, the *GBD* algorithm is implemented in LINGO 18.0 software on some small and large sized networks, where the convergence of the algorithm and the

Table 3. The data of the network in Figure (2)

Nodes	1	2	3	4	5	6	7	8	9
e	1600	1000	1000	1800	1000	1000	1600	1600	1400
w	70	110	80	70	80	100	60	80	60
a	430	320	370	450	300	480	280	350	400

validity of the solutions to the $TCPM$ model are verified. In addition to the convex combination method, there are also some developed solvers for convex nonlinear programming problems which have been applied to solve the primal problem (4.3) during the examples in this section.

In the first example, a small network with 9 nodes shown in Figure (2) is considered, where $I = J = N$. The free-flow travel costs of links, t_e^0, and their capacities, c_e, are illustrated next to them. A self-loop is assigned to each node with $t_e^0 = 0$ and $c_e = M$ where M is a large number. The other necessary information of the network is represented in Table (3).

In addition, the specified demands of $O - D$ pairs, q_{rs}s, are represented in $O - D$ trip matrix A:

$$A = \begin{bmatrix} 0 & 15 & 3 & 9 & 5 & 7 & 9 & 15 & 2 \\ 32 & 0 & 34 & 18 & 18 & 12 & 32 & 13 & 4 \\ 28 & 14 & 0 & 9 & 15 & 4 & 5 & 33 & 34 \\ 21 & 3 & 9 & 0 & 13 & 29 & 1 & 2 & 6 \\ 23 & 26 & 23 & 16 & 0 & 20 & 11 & 27 & 7 \\ 25 & 7 & 13 & 22 & 28 & 0 & 3 & 33 & 28 \\ 18 & 16 & 16 & 11 & 18 & 18 & 0 & 29 & 28 \\ 23 & 14 & 29 & 19 & 13 & 33 & 31 & 0 & 20 \\ 22 & 21 & 8 & 11 & 17 & 9 & 30 & 7 & 0. \end{bmatrix} \tag{5.1}$$

Using *Eppstein*'s k-shortest path ranking algorithm [14], 117 paths are picked up for all $O - D$ and $C - F$ pairs. The proposed GBD method has been applied to solve the $TCPM$ problem (2.5) for $p = 2, 3$. The number of iterations, $(Iter_{GBD})$, required for the upper and lower bounds to meet together, along with the corresponding elapsed running time in terms of seconds (ET_{GBD}) and the optimal objective function $(OBJF_{GBD})$ have been reported in Table (4). In addition, the selected facilities with the obtained values for x_{ij} are also inserted in each case. To verify the validity of the obtained solutions, the $TCPM$ problem (2.5) has been solved by LINGO 18.0 nonlinear mixed-integer branch and bound solver, where the objective function value $(OBJF_{LINGO})$ and the running time (ET_{LINGO}) are represented in each case. The value of the objective function obtained by the GBD method is equal to that of the $LINGO$ branch and bound solver. The variations of the upper and lower bounds in the GBD algorithm, in case $p = 2$, have been represented in Figure (3). The convergence is achieved in the 24th iteration.

Table 4. The estimated solutions for the 9 node network, given in Figure (2)

The solution x_{ij} of GBD for $p=2$ with the given capacities									
$Iter_{GBD}=24,\ ET_{GBD}=23.28\ \ OBJF_{GBD}=61.25(10^3)$ $ET_{LINGO}=3,\ OBJF_{LINGO}=61.25(10^3)$									
$i\backslash j$	1	2	3	4	5	6	7	8	9
3	70	110	80	0	0	100	0	0	0
8	0	0	0	70	80	0	60	80	60
The solution x_{ij} of GBD for $p=3$ with the given capacities									
$Iter_{GBD}=77,\ ET_{GBD}=62\ \ OBJF_{GBD}=57.32(10^3)$ $ET_{LINGO}=3,\ OBJF_{LINGO}=57.32(10^3)$									
$i\backslash j$	1	2	3	4	5	6	7	8	9
1	70	110	0	70	0	0	0	0	0
6	0	0	80	0	0	100	0	0	0
8	0	0	0	0	80	0	60	80	60

The solution x_{ij} of GBD for $p=2$ with the reduced capacities									
$Iter_{GBD}=23,\ ET_{GBD}=19\ \ OBJF_{GBD}=36.80(10^4)$ $ET_{LINGO}=3,\ OBJF_{LINGO}=36.80(10^4)$									
$i\backslash j$	1	2	3	4	5	6	7	8	9
4	70	8.56	0	70	80	0	60	80	0
6	0	101.44	80	0	0	100	0	0	60
The solution x_{ij} of GBD for $p=3$ with the reduced capacities									
$Iter_{GBD}=63,\ ET_{GBD}=54\ \ OBJF_{GBD}=36.26(10^4)$ $ET_{LINGO}=2,\ OBJF_{LINGO}=36.26(10^4)$									
$i\backslash j$	1	2	3	4	5	6	7	8	9
2	70	110	14.47	0	0	0	0	0	0
6	0	0	65.53	0	74.32	100	0	0	60
7	0	0	0	70	5.68	0	60	80	0

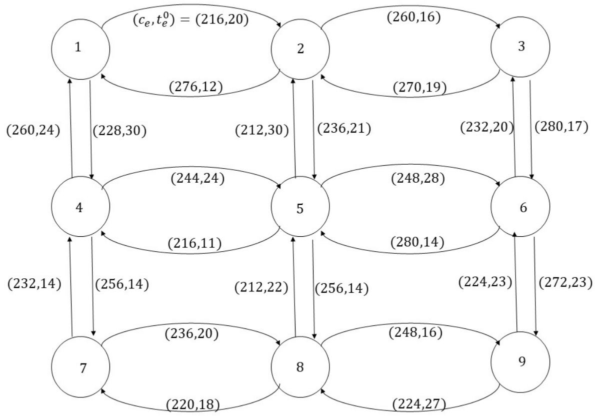

Figure 2. The grid network with 9 nodes

In Table (4), node 8 is selected as a facility in both cases $p = 2, 3$. To perceive the effect of congestion on locating facilities and assigning demands, we reduce the capacities of links entering node 8, except its self-loop, in both cases $p = 2, 3$ to $c_e = 20$, and solve the problem via both the GBD algorithm and $LINGO$ solver. The obtained solutions are shown in Table (4). In both cases $p = 2, 3$, some node is substituted for node 8. Simultaneously, the model tends to impose as least traffic flow as possible to the links entering node 8 to minimize the cost of traffic congestion. However, considering the $O-D$ trip demands $x_{r,8}$ in matrix A, assigning some flows to the links entering node 8 is unavoidable. Nevertheless, the links entering node 8 endure much less traffic flow in the case of reduced capacities compared with the given capacities, see Table (5). Figures (4) and (5) illustrate the selected facilities and the used paths by customers to get service from facilities in both cases with the given and reduced capacities for $p = 2$. To avoid confusion and complexity, the flows related to $O-D$ trips are ignored in the figures.

If the effect of $O-D$ trip demands on traffic congestion is not considered i.e., $q_{rs} = 0$ for all $O-D$ pairs (r, s) and the link capacities decrease, then nodes 4 and

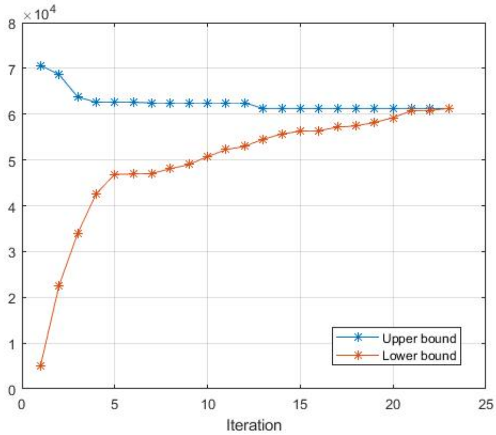

Figure 3. The variations of the upper and lower bounds

Table 5. Link traffic counts for 9-node network

$p = 2$, $v_{e=(i,j)}$ for given capacities

$i \backslash j$	1	2	3	4	5	6	7	8	9
1	0	161	–	18	–	–	–	–	–
2	73.07	0	261	–	107	–	–	–	–
3	–	73.13	80	–	–	119	–	–	–
4	162.93	–	–	78	0	–	139	–	–
5	–	49.93	–	183.93	0	90	–	214	–
6	–	–	204.13	–	219.87	0	–	–	62
7	–	–	–	139	–	–	0	195	–
8	–	–	–	–	27	–	103	80	129
9	–	–	–	–	–	150	–	97	0

$p = 2$, $v_{e=(i,j)}$ for reduced capacities

$i \backslash j$	1	2	3	4	5	6	7	8	9
1	0	76	–	103	–	–	–	–	–
2	130	0	197.85	–	130.3	–	–	–	–
3	–	152.74	0	–	–	270.7	–	–	–
4	106	–	–	70	78	–	84	–	–
5	–	72.41	–	249.55	0	90	–	113	–
6	–	–	138.59	–	140.26	100	–	–	68
7	–	–	–	245	–	–	0	80	–
8	–	–	–	–	70.41	–	149	0	113.59
9	–	–	–	–	–	200.56	–	37	0

$p = 3$, $v_{e=(i,j)}$ for given capacities

$i \backslash j$	1	2	3	4	5	6	7	8	9
1	70	91	–	18	–	–	–	–	–
2	217	0	104	–	107	–	–	–	–
3	–	117	0	–	–	199	–	–	–
4	199	–	–	0	60.32	–	69	–	–
5	–	63	–	150	0	49.32	–	214	–
6	–	–	125	–	176	100	–	–	62
7	–	–	–	121.32	–	–	0	142.68	–
8	–	–	–	–	27	–	103	80	146.68
9	–	–	–	–	–	167.68	–	97	0

$p = 3$, $v_{e=(i,j)}$ for reduced capacities

$i \backslash j$	1	2	3	4	5	6	7	8	9
1	0	146	–	33	–	–	–	–	–
2	107	110	128.15	–	86	–	–	–	–
3	–	131.47	0	–	–	190.53	–	–	–
4	129	–	–	0	78	–	159.68	–	–
5	–	81.15	–	175.86	0	141.32	–	113	–
6	–	–	106.85	–	176	100	–	–	68
7	–	–	–	118.83	–	–	60	80	–
8	–	–	–	–	65.32	–	162.83	0	104.85
9	–	–	–	–	–	191.85	–	37	0

6 are chosen as the 2-median solution of the $TCPM$ model (2.5) with the objective function value $15.59(10^3)$. For $p = 3$, the selected nodes for facilities are $2, 6, 7$ with the objective function value $12.23(10^3)$. In both cases, node 8 is not selected as a facility, and the traffic flow on its entering links is zero. Indeed, the model puts these links aside from being applied for transportation in order to reduce the cost of traffic congestion. However, if the demand of node 8 was much greater than that of the other nodes, the $TCPM$ model might locate one facility at node 8 to provide its demand. To take the situation, let $w_8 = 200$, then for $p = 2$, facilities would be located at

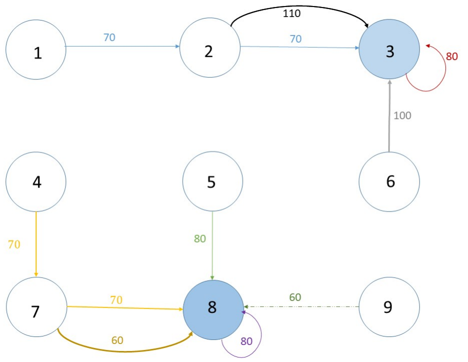

Figure 4. The selected paths and facilities for $p = 2$ with given capacities

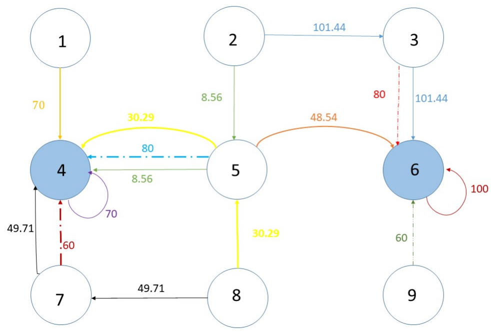

Figure 5. The selected paths and facilities for $p = 2$ with reduced capacities

nodes $1, 9$ and for $p = 3$, the selected nodes would be $1, 3, 8$. Indeed, in case $p = 2$, one facility is absorbed close to node 8, while for $p = 3$, the model locates one facility at node 8, which is mainly devoted for providing its own demand, while the other two facilities are assigned to the rest. Selecting node 8 causes its entering links to endure marginal flows, almost equal to 20. Although the major part of the capacity of facility 8 is devoted to serve its own demand, it also has a small share in providing some other nodes' demands to minimize the traffic congestion and burden imposed on facilities $1, 3$. Generally speaking, large capacities of entering links and small capacities of exiting links corresponding to some node will presumably increase the chance of establishing a facility at this node. However, the effect of clients' demands and capacities of candidate nodes as well as the existing $O - D$ trips should not be neglected.

Next, the influence of such parameters as link capacities and distances between nodes, as well as their demands, on selecting facilities and allocating customers, has been investigated more precisely. We have considered equal values for e_i, w_j, t_e^0, c_e, a_i

and q_{rs} for all $i \in I$, $j \in J$, $e \in E$ and $(r, s) \in ODset$. Some changes have been created in different parameters, then the optimal locations of 2 facilities have been determined in each case.

- Case 1: $e_i = 1000$, $w_j = 60$, $t_e^0 = 10$, $c_e = 200$, $a_i = 300$ and $q_{rs} = 50$ for all $j \in J$, $i \in I$, $e \in E$ and $(r, s) \in ODset$ (for self-loops $t_e^0 = 0$, $c_e = M$). As shown in Table (6), the solution of the $TCPM$ problem (2.5) is $2, 8$ with $OBJF = 16.36(10^4)$. Indeed, the facilities are located regarding the distances between nodes. However, due to the effect of congestion, the solution is still different from that of the classic p-median problem.

- Case 2: Reducing the capacities of links entering node 2 to $c_e = 20$. The capacities of links leading to node 2 are decreased to 20, while the other parameters remain unchanged as in case 1. Due to the small capacities, the congestion on links entering node 2 increases. Therefore, as seen in Table (6), the locations of facilities change to nodes $4, 6$. Although the demands are allocated in such a way to devote as fewer traffic counts as possible to the links entering node 2, the low-capacity links would still endure some flows related to the $O - D$ trips $x_{r,2}$ for different origins r; therefore, the objective function value increases to $19.63(10^6)$.

  If facilities are located again at nodes $2, 8$, the objective function will increase to $61.25(10^6)$, of which more than 99.7% is related to the congestion costs of links entering node 2. To sum up, the traffic congestion imposed on the network to provide the demands by using a facility has a considerable effect on the efficiency of this facility. Facilities are to be located in nodes having as good access as possible, while the other parameters should also be considered.

- Case 3: Increasing the demand of node 3 to $w_3 = 120$. We increase the demand of node 3 to 120, while the other parameters remain unchanged as case 1. Here, facilities are located at nodes $3, 4$, see Table (6). Indeed, placing one median at node 3 and assigning its demand to its own, will decrease the traffic congestion related to providing service for this node. Also, the demands of other nodes are generally allocated to their nearest facility provided that the capacities of the nearest facility and its connected links are large enough.

- Case 4: Reducing the capacities of links entering node 3 to $c_e = 20$ and increasing its demand to $w_3 = 120$. In this case, the model avoids selecting node 3 as a median, see Table (6); because the small capacities of its entering links will generate excess traffic congestion and prevent users from getting service. However, its considerable demand absorbs one median to its proximity

In the second example, we consider Siouxfalls network represented in Figure (6) in both urban and network maps, where such necessary information as capacities and free-flow travel costs of links as well as the $O - D$ trip demands have been provided in the URL address $https://github.com/bstabler/TransportationNetworks$. The

Table 6. **The solutions to the traffic-based capacitated 2-median problem**

Case 1	1	2	3	4	5	6	7	8	9	$OBJF$	Case 2	2	2	3	4	5	6	7	8	9	$OBJF$
2	60	60	60	27.94	34.92	27.13	0	0	0	$16.36(10^4)$	4	60	30	0	60	30	0	60	30	0	$19.63(10^6)$
8	0	0	0	32.06	25.08	32.87	60	60	60		6	0	30	60	0	30	60	0	30	60	
Case 3	1	2	3	4	5	6	7	8	9	$OBJF$	Case 4	1	2	3	4	5	6	7	8	9	$OBJF$
3	0	60	120	0	0	60	0	0	60	$16.52(10^4)$	4	60	60	0	60	56.89	0	60	3.11	0	$61.85(10^5)$
4	60	0	0	60	60	0	60	60	0		6	0	0	120	0	3.11	60	60	56.89	60	

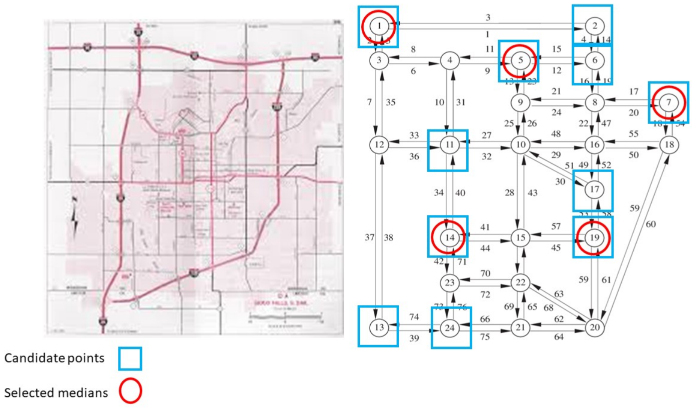

Figure 6. **Siouxfalls network map**

network has 24 nodes and 100 links (regarding self-loops), where all nodes are considered to have some demands. Since there is not any data available for the service demands of Siouxfalls nodes, we have randomly taken some values from the interval $[150, 550]$. In addition, 12 candidate nodes, represented by the blue squares in Figure (6), have been randomly selected for establishing p facilities where the values of fixed set-up costs and capacities of candidate nodes have been taken from the intervals $[3000, 5500]$ and $[1500, 3500]$, respectively. Also, one self-loop link with a large capacity and zero travel cost is associated to each candidate node. Applying *Eppstein*'s k-shortest path ranking algorithm [14], 940 paths are picked up for all $O-D$ and $C-F$ pairs.

The problem is solved for $p = 5$ by applying the GBD algorithm by which, after 96 iterations, taking 1017 seconds, the convergence of the upper and lower bounds is achieved. The selected facilities are located at nodes $1, 5, 7, 14, 19$ depicted in Figure (6) in the red circles. As it is seen, two facilities $1, 5$ are picked up from the top part of the network. Then, to verify the effect of congestion imposed by the existing $O-D$ flows on determining the location of facilities, we have multiplied the amounts of entering $O-D$ flows to the top part of the city by 3; i.e. $x_{r,s} = 3x_{r,s}$ for $r = 1, ..., 24$ and $s = 1, 2, 3, 4, 5, 6$. New $O-D$ demands result in facilities at nodes $6, 7, 11, 14, 19$. Indeed, the model avoids locating facilities inside the congested areas; instead, one facility is located on the border of this area to service the upper nodes while the others

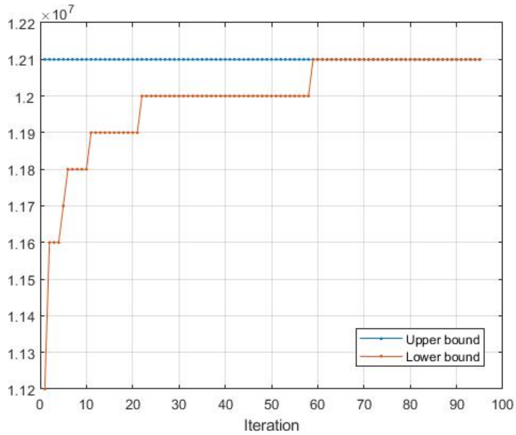

Figure 7. **The variations of the upper and lower bound for $p = 4$**

Table 7. **The results for Siouxfalls network**

p	5		6		7	
	$OPTOBJ(10^6)$	ET_{OPT}	$OPTOBJ(10^6)$	ET_{OPT}	$OPTOBJ(10^6)$	ET_{OPT}
	12.09	1049	12.05	1114	12.04	2430
N_{iter}	20	30	20	30	20	30
ET	273	359	251	382	400	428
$UB(10^6)$	12.11	12.11	12.06	12.06	12.06	12.05
$LB(10^6)$	11.95	11.99	11.94	11.97	11.92	11.95
RD	0.013	0.010	0.010	0.007	0.012	0.008
RE	0.002	0.002	0.008	0.008	0.002	0.001

are set on the center or bottom part of the city.

The variations of the upper and lower bounds, without multiplying the $O-D$ trip values, are represented in Figure (7). Noting Figure (7), although it takes 96 iterations for the upper and lower bounds to meet each other, their relative difference is subtle after 60 iterations. In addition, the upper bound remains unchanged during all iterations, while the lower bound increases. Indeed, conducting different examples, it has been realized that the upper bound usually meets its optimal value much sooner than the lower bound. Therefore, we suggest stopping the algorithm when no improvement is achieved in the upper bound during a specified number of successive iterations. Using this condition, the $TCPM$ problem for $p = 5, 6, 7$ is solved for which the GBD algorithm would be stopped if no improvement is observed in the upper bound after N_{iter} successive iterations. Then, the elapsed running time (ET) and the values of the upper and lower bounds (UB, LB) with their relative difference (RD) obtained for different values of $N_{iter} = 20, 30$ along with the optimal objective function value ($OPTOBJ$) (calculated by solving the problem with no iteration limit), its corresponding running time (ET_{OPT}), and the relative error between $OPTOBJ$ and UB represented by RE, have been inserted in Table (7).

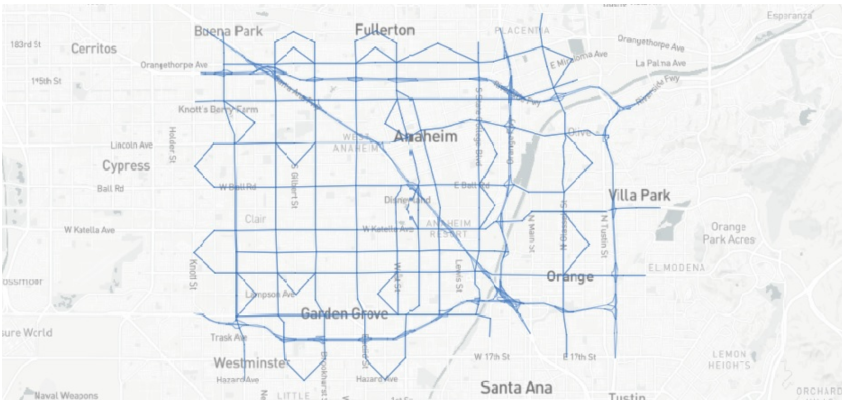

Figure 8. **Anaheim network**

As it is realized by Table (7), setting a reasonable limit for successive iterations with no progress, for example $N_{iter} = 20$, results in a close approximation to the optimal objective value in a short running time, almost one-fifth of the time required for the global convergence. This observation increases the computational efficiency of the algorithm by picking the upper bound as a fair estimate of the optimal solution. Besides, noting Table (7), it is perceived that the optimal value of the objective function decreases by increasing the number of facilities, because the traffic congestion, as well as the total travel cost, would be reduced by assigning customers to closer facilities.

In the last example, Anaheim network with 416 nodes, 914 links, and 38 zones, represented in Figure (8) is considered whose topology, including the values of c_e and t_e^0, is available at the URL address $https : //github.com/bstabler /TransportationNetworks$. The zones are assumed to be the origins and destinations which generate or absorb trips where the demands of all 38×38 $O-D$ pairs are available in the mentioned address. We have selected 25 candidate nodes to establish facilities whose capacities have been taken from the interval $[1300, 2800]$. The fixed establishment costs have been assumed equal to zero to evaluate the effect of link travel costs on selecting facilities. Also, one self-loop link with a large capacity and zero travel cost is associated to each candidate node which causes the number of links increase to 1330. In addition, the demand points are located at the network zones, for which the values of w_j belong to $[150, 650]$. Applying *Eppstein*'s k-shortest path ranking algorithm [14], 1896 and 2850 paths are picked up for $C-F$ and $O-D$ pairs, respectively.

To realize the effect of the number of established facilities as well as link capacities on toll values, we have solved the problem for different values of p and capacities of links. The link tolls are estimated by $\gamma_e = \tau_e(v_e^{opt})v_e^{opt}$ where v_e^{opt}s are the optimal traffic counts of links. The algorithm has been implemented for $p = 7, 10, 13$ and link capacities $c_e^{\mu} = \mu c_e$ with $\mu = 0.5, 1, 2$ considering $N_{iter} = 20, 30$ (maximum number of

Table 8. **The estimated solutions of Anaheim Network**

$N_{iter} = 20$	$p = 7$			$p = 10$			$p = 13$		
	$\beta = 0.5$	$\beta = 1$	$\beta = 2$	$\beta = 0.5$	$\beta = 1$	$\beta = 2$	$\beta = 0.5$	$\beta = 1$	$\beta = 2$
ET	1689	1685	1680	2042	1667	1384	1126	1440	988
$UB(10^5)$	42.29	14.57	12.70	41.47	14.28	12.47	41.43	14.21	12.45
$LB(10^5)$	42.27	14.53	12.69	41.32	14.13	12.37	41.35	14.14	12.37
$RD(10^{-3})$	0.4	3	0.8	3.6	10.5	8	1.9	4.9	6.4
AT	303.64	19.39	1.46	297.62	19.04	1.37	297.37	18.89	1.36
$N_{iter} = 30$	$p = 7$			$p = 10$			$p = 13$		
	$\beta = 0.5$	$\beta = 1$	$\beta = 2$	$\beta = 0.5$	$\beta = 1$	$\beta = 2$	$\beta = 0.5$	$\beta = 1$	$\beta = 2$
ET	1896	2012	1555	2100	2359	1877	7901	4009	1498
$UB(10^5)$	42.29	14.57	12.70	41.47	14.28	12.47	41.43	14.21	12.43
$LB(10^5)$	42.29	14.57	12.70	41.32	14.15	12.38	41.35	14.15	12.37
RD	0	0	0	3.6	9.1	7.2	1.9	4	4.8
AT	303.64	19.39	1.46	297.62	19.04	1.37	297.37	18.89	1.36

successive iterations with no improvement in the upper bound). The average amount of link tolls (AT) has been calculated by finding the optimal solution, and inserted in Table (8). Also, the upper and lower bounds (UB, LB) and their relative difference (RD), along with the elapsed running time (ET) have been reported in Table (8).

By increasing the number of established facilities or capacities of links, the average toll decreases. Indeed, if the system planner attempts to impose less financial burden on users, he/she must establish more facilities in the network, or instead, he/she should improve the link capacities by some development plans. Both measures would also cause a decline in the objective function value, provided that the establishment costs are zero. In addition, by increasing N_{iter}, the relative difference between the upper and lower bounds gets smaller; however the running time increases.

6. Summary and conclusions

In this paper, considering the effect of link traffic congestion on travel cost, we introduced a traffic-based model to the capacitated p-median problem. Indeed, a combination of the system optimization model with the p-median problem, having fixed set-up costs for establishing facilities at nodes, was applied. It was assumed that the clients' demands were to be provided through different paths to facilities. Besides, the effect of the existing $O - D$ trip demands was simultaneously taken into consideration. The purpose was to minimize the total network establishment and transportation cost. The generalized Benders decomposition method was applied to decompose the problem into some easier sub-problems and the solution was iteratively updated to minimize the difference between the upper and lower bounds. Also, the convergence of the proposed method in a finite number of iterations was addressed through a theorem. Finally, some numerical examples have been analyzed to illustrate the efficiency of the proposed model and the solution approach.

Acknowledgements: The first and the second authors were supported in part by the Iranian National Science Foundation (INSF) under Grant Number 99012804.

Conflict of Interest: The authors declare that they have no conflict of interest.

Data Availability: Data sharing is not applicable to this article as no data sets were generated or analyzed during the current study.

References

[1] M. Abareshi and M. Zaferanieh, *A bi-level capacitated p-median facility location problem with the most likely allocation solution*, Transportation Research Part B: Methodological **123** (2019), 1–20.
https://doi.org/10.1016/j.trb.2019.03.013.

[2] M. Abareshi, M. Zaferanieh, and B. Keramati, *Path flow estimator in an entropy model using a nonlinear l-shaped algorithm*, Networks and Spatial Economics **17** (2017), no. 1, 293–315.
https://doi.org/10.1007/s11067-016-9327-9.

[3] R. Aboolian, O. Berman, and M. Karimi, *Probabilistic set covering location problem in congested networks*, Transp. Sci. **56** (2022), no. 2, 528–542.
https://doi.org/10.1287/trsc.2021.1096.

[4] T. Abrahamsson, *Estimation of origin-destination matrices using traffic counts-a literature survey*, 1998.

[5] M.J. Bagajewicz and V. Manousiouthakis, *On the generalized benders decomposition*, Comput. Chem. Eng. **15** (1991), no. 10, 691–700.
https://doi.org/10.1016/0098-1354(91)85015-M.

[6] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*, John wiley & sons, 2006.

[7] M.J. Beckmann, C.B. McGuire, and C.B. Winsten, *Studies in the economics of transportation*, Yale University Press, New Haven, 1956.

[8] C. Beltran, C. Tadonki, and J.P. Vial, *Solving the p-median problem with a semi-Lagrangian relaxation*, Comput. Optim. Appl. **35** (2006), no. 2, 239–260.
https://doi.org/10.1007/s10589-006-6513-6.

[9] J.F. Benders, *Partitioning procedures for solving mixed-variables programming problems*, Numer. Math. **4** (1962), no. 1, 238–252.
https://doi.org/10.1007/BF01386316.

[10] P. Bergendorff, D.W. Hearn, and M.V. Ramana, *Congestion Toll pricing of Traffic Networks*, Network Optimization (Berlin, Heidelberg) (P.M. Pardalos, D.W. Hearn, and W.W. Hager, eds.), Springer Berlin Heidelberg, 1997, pp. 51–71
https://doi.org/10.1007/978-3-642-59179-2_4.

[11] C. Bütün, S. Petrovic, and L. Muyldermans, *The capacitated directed cycle hub location and routing problem under congestion*, European J. Oper. Res. **292** (2021),

no. 2, 714–734.
https://doi.org/10.1016/j.ejor.2020.11.021.
[12] J.A. Díaz and E. Fernandez, *Hybrid scatter search and path relinking for the capacitated p-median problem*, European J. Oper. Res. **169** (2006), no. 2, 570–585.
https://doi.org/10.1016/j.ejor.2004.08.016.
[13] J. Dupuit, *De la measure de l'utilité des travaux publiques*, Annales des Ponts et Chaussées **8** (1844), 255–283.
[14] D. Eppstein, *Finding the k shortest paths*, SIAM J. comput. **28** (1998), no. 2, 652–673.
https://doi.org/10.1137/S0097539795290477.
[15] K. Fleszar and K.S. Hindi, *An effective VNS for the capacitated p-median problem*, European J. Oper. Res. **191** (2008), no. 3, 612–622.
https://doi.org/10.1016/j.ejor.2006.12.055.
[16] C.A. Floudas, A. Aggarwal, and A.R. Ciric, *Global optimum search for nonconvex NLP and MINLP problems*, Comput. Chem. Eng. **13** (1989), no. 10, 1117–1132.
https://doi.org/10.1016/0098-1354(89)87016-4.
[17] M. Frank and P. Wolfe, *An algorithm for quadratic programming*, Naval research logistics quarterly **3** (1956), no. 1-2, 95–110.
[18] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness*, Freeman, 1979.
[19] A.M. Geoffrion, *Duality in nonlinear programming: a simplified applications-oriented development*, SIAM review **13** (1971), no. 1, 1–37.
https://doi.org/10.1137/1013001.
[20] ———, *Generalized benders decomposition*, J. Optim. Theory Appl. **10** (1972), no. 4, 237–260.
https://doi.org/10.1007/BF00934810.
[21] M. Golabi, S.M. Shavarani, and L. Idoumghar, *A congested capacitated location problem with continuous network demand*, RAIRO Oper. Res. **56** (2022), no. 5, 3561–3579.
https://doi.org/10.1051/ro/2022167.
[22] S.L. Hakimi, *Optimum distribution of switching centers in a communication network and some related graph theoretic problems*, Oper. Res. **13** (1965), no. 3, 462–475.
https://doi.org/10.1287/opre.13.3.462.
[23] O. Kariv and S. Hakimi, *An algorithmic approach to network location problems. II: the p-medians*, SIAM J. Appl. Math **37** (1979), 539–560.
https://doi.org/10.1137/0137041.
[24] L.A.N. Lorena and E.L.F. Senne, *A column generation approach to capacitated p-median problems*, Comput. Oper. Res. **31** (2004), no. 6, 863–876.
https://doi.org/10.1016/S0305-0548(03)00039-X.
[25] V. Maniezzo, A. Mingozzi, and R. Baldacci, *A bionomic approach to the capacitated p-median problem*, J. Heuristics **4** (1998), no. 3, 263–280.
https://doi.org/10.1023/A:1009665717611.

[26] M.D. Meyer, *Transportation Planning Handbook*, Wiley, 2016.
[27] A.W. Neebe, *Branch and bound algorithm for the p-median transportation problem*, J. Oper. Res. Soc. **29** (1978), no. 10, 989–995.
https://doi.org/10.1057/jors.1978.212.
[28] M. Patriksson, *The Traffic Assignment Problem: Models and Methods*, Dover Publications, 2015.
[29] A.C. Pigou, *The Economics of Welfare*, Creative Media Partners, LLC, 2022.
[30] J. Reese, *Solution methods for the p-median problem: An annotated bibliography*, Networks **48** (2006), no. 3, 125–142.
https://doi.org/10.1002/net.20128.
[31] C.S. ReVelle and R.W. Swain, *Central facilities location*, Geographical Analysis **2** (1970), no. 1, 30–42.
https://doi.org/10.1111/j.1538-4632.1970.tb00142.x.
[32] H. Shamsipoor, M.A. Sandidzadeh, and M. Yaghini, *Solving capacitated p-median problem by a new structure of neural network*, Int. J. Indus. Eng. **19** (2012), no. 8, 305–319.
[33] Y. Sheffi, *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall, New Jersey, 1984.
[34] S. Shiripour, N. Mahdavi-Amiri, and I. Mahdavi, *A transportation network model with intelligent probabilistic travel times and two hybrid algorithms*, Transp. Lett. **9** (2017), no. 2, 90–122.
https://doi.org/10.1080/19427867.2016.1187893.
[35] A. Tamir, *An $O(pn^2)$ algorithm for the p-median and related problems on tree graphs*, Oper. Res. Lett. **19** (1996), no. 2, 59–64.
https://doi.org/10.1016/0167-6377(96)00021-1.
[36] J.G. Wardrop, *Some theoretical aspects of road traffic research*, Proceedings of the Institute of Civil Engineers **1** (1952), no. 2, 325–362.
[37] M. Zaferanieh, *The most probable allocation solution for the p-median problem*, Iranian Journal of Numerical Analysis and Optimization **10** (2020), no. 2, 155–176
https://doi.org/10.22067/ijnao.v10i2.84892.
[38] M. Zaferanieh, M. Abareshi, and J. Fathali, *The minimum information approach to the uncapacitated p-median facility location problem*, Transp. Lett. **14** (2022), no. 3, 307–316.
https://doi.org/10.1080/19427867.2020.1864595.
[39] M. Zaferanieh, M. Abareshi, and M. Jafarzadeh, *A bi-level p-facility network design problem in the presence of congestion*, Comput. Indus. Eng. **176** (2023), Article ID: 109010.
https://doi.org/10.1016/j.cie.2023.109010.
[40] M. Zaferanieh, M. Sadra, and T. Basirat, *P-facility capacitated location problem with customer-equilibrium decisions: a recreational case study in mazandaran province*, J. Modelling in Management (2024), In press.
https://doi.org/10.1108/JM2-08-2023-0194.