Applied Soft Computing 33 (2015) 24–36

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

A hybridization of an evolutionary algorithm and a parallel branch and bound for solving the capacitated single allocation hub location problem☆

Predrag Stanojević, Miroslav Marić, Zorica Stanimirović*

Faculty of Mathematics, University of Belgrade, Studentski trg 16/IV, 11 000 Belgrade, Serbia

ARTICLE INFO

Article history:
Received 28 October 2013
Received in revised form 20 October 2014
Accepted 6 April 2015
Available online 17 April 2015

Keywords:
Metaheuristic
Hybrid algorithms
Evolutionary algorithm
Branch-and-bound
Network design
Hub location problems

ABSTRACT

In this study, we propose a hybrid optimization method, consisting of an evolutionary algorithm (EA) and a branch-and-bound method (BnB) for solving the capacitated single allocation hub location problem (CSAHLP). The EA is designed to explore the solution space and to select promising configurations of hubs (the location part of the problem). Hub configurations produced by the EA are further passed to the BnB search, which works with fixed hubs and allocates the non-hub nodes to located hubs (the allocation part of the problem). The BnB method is implemented using parallelization techniques, which results in short running times. The proposed hybrid algorithm, named EA-BnB, has been tested on the standard Australia Post (AP) hub data sets with up to 300 nodes. The results demonstrate the superiority of our hybrid approach over existing heuristic approaches from the existing literature. The EA-BnB method has reached all the known optimal solutions for AP hub data set and found new, significantly better, solutions on three AP instances with 100 and 200 nodes. Furthermore, the extreme efficiency of the implementation of this hybrid algorithm resulted in short running times, even for the largest AP test instances.

1. Introduction

Hub location problems arise from transportation and telecommunication networks when it is not desirable to directly transport goods, passengers or data between origin–destination pairs, due to extremely high transportation costs. As an alternative, a hub network is used, where hubs act as collection, consolidation, transfer and distribution points. The advantage of exploiting a hub network are lower transportation costs between the hubs, which leads to reductions of overall transportation costs in the network. Origin and destination nodes can be connected to one or more hubs, depending on whether the design constraints allow single or multiple allocation. It is usually assumed that the underlying hub network is fully connected, while non-hub nodes are not necessarily connected to each other. Furthermore, all origin–destination flow has to be routed via at least one hub.

Hub location problems usually involve two decision making tasks: choosing which hubs to establish from the given set of potential hubs and the allocation of non-hub nodes to established hubs. Various constraints and objective functions may be assumed. The most common hub location problems are those with the center and median objectives. In the hub median problems, the objective is to minimize the total transportation costs in the network, which is important in transportation systems, such as air cargo and passenger transport, postal and other delivery systems, etc. However, in the case of excessively large or expensive origin–destination distances, the objective of the median type may lead to unsatisfactory results. In these cases, hub center problems represent a better model, minimizing the maximum distance or cost between origin–destination pairs. Hub center problems are mostly applied in designing fast delivery systems (DHL, Fedex, LightSpeed Express, etc.), which are used for urgent deliveries and transportation of perishable or time sensitive items. Various additional constraints may be imposed, such as fixed number of hubs to be located, limited capacities of both hub and non-hub nodes, capacity constraints on links in the network, fixed costs for establishing hubs or hub network, etc. A detailed review of hub location problems and their applications may be found in [3,6].

In this paper, we consider a variant of the hub location problem, in the literature known as the capacitated single allocation hub location problem (CSAHLP). In this problem there is capacity restriction on the incoming flow of each potential hub node. The

☆ This research was partially supported by Serbian Ministry of Education, Science and Technological Development under the grant nos. 174010 and 47017.

* Corresponding author. Tel.: +381 112027801; fax: +381 112630151.
E-mail addresses: djape@sbb.rs (P. Stanojević), maricm@matf.bg.ac.rs (M. Marić), zoricast@matf.bg.ac.rs (Z. Stanimirović).

http://dx.doi.org/10.1016/j.asoc.2015.04.018
1568-4946/

number of hubs to be installed is not fixed in advance, but installing a hub at some location assumes certain fixed costs. We want to choose locations for installing hubs and to allocate each non-hub node to exactly one, previously installed hub, in such a way that the sum of transportation costs between origin–destination pairs and the costs for establishing hubs is minimized. The CSAHLP is an NP-hard optimization problem, since its uncapacitated variant (USAHLP) is known to be NP-hard, even when the set of hubs is fixed [18].

The CSAHLP was not so extensively studied in the literature, compared to its uncapacitated version – the USAHLP. There is a plethora of papers dealing with the uncapacitated single allocation hub location problem. Up to now several formulations of the USAHLP were proposed in the literature [2,14,18], and several exact and heuristic methods were developed for solving the USAHLP, see [1,2,21,25]. We refer to some recent studies dealing with the USAHLP, see [7,13,16,20].

In the literature one can find several papers that consider the CSAHLP and its appropriate solution methods. Campbel [5] presents the first mixed integer linear programming (MILP) formulation for the problem. Ernst and Krishnamoorthy in [11] extend the formulation of the uncapacitated version of the problem from [22] to the capacitated case. The authors in [11] present another MILP formulation of the problem by involving an additional set of constraints. This formulation was later corrected by Correia et al. in [9] by adding a missing constraint set. Two heuristic algorithms for solving the CSAHLP are proposed in [11], based on simulated annealing (SA) and random descent (RDH) approaches. The upper bounds obtained by the heuristics are used in an LP-based branch- and bound method, which provides optimal solutions for small and medium size AP problem instances with $n \leq 50$ nodes. For realistic sized AP problems $n=100$, 200 that could not be solved exactly, the proposed RDH and SA heuristics provided solutions in a reasonable amount of computer time.

Contreras et al. [8] present a Lagrangian relaxation (LR) enhanced with reduction tests, which exploits the structure of the problem and decomposes it into smaller subproblems that can be solved efficiently. The authors present optimal solutions obtained by LR approach for small and medium size AP instances with $n \leq 50$ nodes, and tight upper and lower bounds in the case of newly generated instances of larger dimensions. A heuristic approach is employed to obtain good quality feasible solutions for the tested instances.

Stanimirović in [23] proposed a genetic algorithm (GA) approach for solving the CSAHLP. The GA from [23] proved to be efficient on small and medium size AP test instances with $n \leq 50$ nodes and had a similar performance on larger AP instances with $n=100, 200$ nodes as the RDH and SA heuristics. In the study by Randall [19], four variations of the ant colony metaheuristic (ACO) are proposed as a solution method for the CSAHLP. The developed ACO approaches use different learning mechanisms for determining the location of hubs and the assignment of non-hub nodes to hubs. The authors investigate the effects of the solution component assignment order, and the form of local search heuristics through the set of computational experiments on small and medium size AP instances with $n \leq 50$ nodes.

A variant of the CSAHLP is studied in paper [15], where a capacity on the flow that transverses each hub is assumed, and a branch-and-cut algorithm is proposed for solving this problem. Costa et al. [10] propose a bi-objective approach to the CSAHLP. Instead of using capacity constraints to limit the amount of incoming flow in hubs, the authors introduce a second objective function to the model that tries to minimize the time to process the flow entering the hubs. For an overview of the existing literature on the CSAHLP and related problems, we refer the reader to [3].

2. Mathematical model

We use a revised formulation of the CSAHLP from [9], including the missing cuts, the lack of which could lead to infeasible solutions, as demonstrated by Correia et al. in [9]. A network I of n distinct nodes is given, with a matrix C_{ij} of transportation cost per unit flow between any two nodes i and j from the network I. In general, this matrix represents any abstract value corresponding to an ordered pair of two nodes. However, in practice, it usually depends on the distance between i and j. Therefore, for a given node k we will label a node l, $l \neq k$ as *the nearest node to node* k, if it has the lowest value of transportation cost from the node k, i.e., $C_{kl} = \min_{j, j \neq k} C_{kj}$.

The amount of flow from an origin i to destination j is given by the flow matrix W_{ij}. This matrix is not necessarily symmetric, and W_{ii} may be greater than 0. The cost of establishing a hub is associated with every node k and is given by a constant f_k. Similarly, the collection capacity of each potential hub k is given by G_k. All flow from an origin i to a destination j has to be routed via some hub nodes k and l, respectively, where $k, l \in H$ and $H \subseteq I$ is a set of established hubs. Therefore, direct transportation between non hub nodes is not allowed. The costs of collection (from origin to hub), transfer (from hub to hub), and distribution (from hub to destination) are given by the parameters χ, α and δ, respectively. Hence, the total transportation cost (per unit flow) from an origin i to a destination j, via hubs k and l, is equal to $\chi C_{ik} + \alpha C_{kl} + \delta C_{lj}$. Since the transfer between the hubs has lower cost compared to collection and distribution, it is assumed that $\chi, \delta > \alpha$.

A binary decision variable Z_{ij} takes the value of 1 if node i is allocated to a hub node $j \in H$, and 0 otherwise. Hubs are always allocated to themselves, hence, if $Z_{kk} = 1$ then and only then k is a hub. $Z_{kk} = 1 \Leftrightarrow k \in H$. No direct flow between non-hub nodes i and j is allowed, so we introduce continuous non-negative variables Y^i_{kl} which represent the amount of flow originated from node i, collected at hub k and distributed via hub l. The triangle-inequality holds for the transportation costs between nodes, so the flow will ever travel via at most two hubs. Finally, O_i and D_j represent the amount of flow which departs from $i \in I$ and the amount of flow that is distributed to node $j \in I$, respectively, i.e., $O_i = \sum_{j \in I} W_{ij}$ and $D_j = \sum_{i \in I} W_{ij}$.

Using the notation given above, the CSAHLP is formulated as follows [9]:

$$\min \sum_{i \in I} \sum_{k \in I} C_{ik} Z_{ik} (\chi O_i + \delta D_i) + \sum_{i \in I} \sum_{k \in I} \sum_{l \in I} \alpha C_{kl} Y^i_{kl} + \sum_{k \in I} f_k Z_{kk} \quad (1)$$

subject to:

$$\sum_{k \in I} Z_{ik} = 1 \quad \text{for every} i \in I \quad (2)$$

$$Z_{ik} \leq Z_{kk} \quad \text{for every} \quad i, k \in I \quad (3)$$

$$\sum_{l \in I} Y^i_{kl} - \sum_{l \in I} Y^i_{lk} = O_i Z_{ik} - \sum_{j \in I} W_{ij} Z_{jk} \quad \text{for every} \quad i, k \in I \quad (4)$$

$$\sum_{i \in I} O_i Z_{ik} \leq G_k Z_{kk} \quad \text{for every} \quad k \in I \quad (5)$$

$$\sum_{l \in I, l \neq k} Y^i_{kl} \leq O_i Z_{ik} \quad \text{for every} \quad i, k \in I \quad (6)$$

$$Y^i_{kl} \geq 0 \quad \text{every} \quad i, k, l \in I \quad (7)$$

$$Z_{ik} \in \{0, 1\} \quad \text{every} \quad i, k \in I. \quad (8)$$