ORIGINAL PAPER

An efficient heuristic for a hub location routing problem

Mustapha Ratli¹ · Dragan Urošević² · Abdessamad Ait El Cadi¹ · Jack Brimberg³ · Nenad Mladenović⁴ · Raca Todosijević^{[1](http://orcid.org/0000-0002-9321-3464)} ®

Received: 31 May 2020 / Accepted: 19 November 2020 / Published online: 28 November 2020 © Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

This paper examines a new model for hub location known as the hub location routing problem. The problem shares similarities with the well studied uncapacitated single allocation p-hub median problem except that the hubs are now connected to each other by a cyclical path (or tour) known as the global route, each cluster of nonhub nodes and assigned hub is also connected by a single tour known as a local route, and the length of the local routes is constrained to a maximum number of nodes. Thus, aside from the normal tasks of hub selection and allocation of nonhub nodes, up to $p + 1$ travelling salesman problems need to be solved. A heuristic based on the general variable neighborhood search framework is proposed here to solve this very complicated problem. The improvement phase of the algorithm uses a sequential variable neighborhood descent with multiple neighborhoods required to suit the complex nature of the problem. A best sequencing of the neighborhoods is established through empirical testing. The perturbation phase known as the shaking procedure also uses a well-structured selection of neighborhoods in order to effectively diversify the search to different regions of the solution space. Extensive computational testing shows that the new heuristic significantly outperforms the state-of-the-art. Out of 912 test instances from the literature, we are able to obtain 691 new best known solutions. Not only are the improvements in objective values quite impressive, but also these new solutions are obtained in a small fraction of the time required by the competing algorithms.

Keywords p -hub · Hub location-routing problem · Heuristics · General variable neighborhood search

1 Introduction

The aim of a transportation/communication network is to facilitate transport/communication between all origin–destination (O–D) pairs. Establishing direct transport/ communication between all O–D pairs may be unreasonable for many reasons espe-

Extended author information available on the last page of the article

cially from the economic (or cost) perspective. As an alternative approach, the idea of hub networks has been proposed. The main feature of a hub network that differentiates it from other types is the existence of special nodes (hubs) designated to enable the connection between O–D pairs. Namely, instead of having a direct path between each pair of nodes, nodes communicate via hubs. For example, if we have an origin node *O* and a destination *D*, a valid path between them may have the form $O - hub_1 - hub_2 - D$. More precisely, the flow from *O* to *D* is first collected at hub_1 then transferred to hub_2 and finally delivered to node D. So, the main principle of hub networks is that direct connection between non-hub nodes is not possible, and instead, all transport/communication is performed via hub nodes. Hub nodes have a role to collect, consolidate and transfer, and finally distribute the flow in a hub network. Note that on a path $O - hub_1 - hub_2 - D$, hub nodes hub_1 and hub_2 may be the same. In this case the same hub is used for collection, consolidation and transfer, and distribution purposes,.

During the past years, many variants of the hub location problem, including those with and without capacity constraints, or with different allocation strategies, topologies and objective functions, have been studied. For more details on the classification, solution approaches and applications of hub location problems, we refer the reader to [\[1](#page-17-0)[,3](#page-17-1)[–6](#page-17-2)[,8](#page-18-0)[–11](#page-18-1)[,14](#page-18-2)[–17](#page-18-3)[,21](#page-18-4)[–23](#page-18-5)[,29](#page-18-6)[–31](#page-18-7)[,34](#page-19-0)[–36](#page-19-1)].

In this paper, we study a relatively new type of hub location problem known as the hub location routing problem. The main feature that differentiates this variant is the existence of local and global routes. A local route starts and ends at a hub that serves one or more non-hub nodes, while the global route is used to interconnect hub nodes, and its main purpose is to transfer the flow between hubs (see Fig. [1\)](#page-1-0).

The problem studied in this paper consists of designating *p* nodes on a given network to serve as hubs, assigning each non-hub node to exactly one hub, connecting non-hub

nodes assigned to the same hub by a route and linking hub nodes by a route so that the total routing cost is minimized. A constraint is placed on the maximum length of the local routes by specifying the maximum number of non-hub nodes allowed on a route. There is neither a capacity limitation imposed on the hubs nor a cost for hub installation. However, as in other hub location problems, a discount factor, $0 \le \alpha \le 1$, is applied when calculating the cost of the global route. The problem studied in this paper is referred to as the *p*-hub location routing problem (pHLRP).

The problem is formally defined on a graph $G = (V, E)$, where V is a given set of nodes and $E = \{(i, j) | i, j \in V\}$ is the set of edges that connect them. A non-negative cost c_e is associated with each edge $e \in E$ representing the routing cost along that edge. A feasible solution of the problem provides *p* nodes from set *V* to serve as hubs. Each of these hubs is the root point of a local route starting and ending at it. Each local route is composed of non-hub nodes, and may contain at most *C* nodes including the hub. If we denote the chosen hubs as $hub_1, hub_2,..., hub_n$ and corresponding local routes as r_1, r_2, \ldots, r_p , where r_i is an ordered list of nodes, then $r_1 \cup r_2 \cup \cdots \cup r_p = V$, and $r_i \cap r_j = \emptyset$, $i \neq j$. In other words, each non-hub node is assigned to exactly one hub. If we suppose that $r_i = \{hub_i, n_1^i, \ldots, n_k^i, hub_i, \}$ then the cost of the route is calculated as

$$
cost(r_i) = c_{(hub_i, n_1^i)} + \sum_{m=1}^{k-1} c(n_m^i, n_{m+1}^i) + c_{(n_k^i, hub_i)}.
$$

If we suppose that the cost of a tour linking the hubs is *cost*(*hub*_*tour*)then the overall cost of the solution defined by hubs hub_1 , hub_2 , ..., hub_p and, corresponding local routes r_1, r_2, \ldots, r_p , is calculated as: $\sum_{i=1}^p cost(r_i) + \alpha cost(hub_tour)$. Note that the length of route r_i , with the above notation is, $length_i = k + 1$. Note also that a route r_i may contain no non-hub nodes at all. In this case the length of the local route is 1, i.e., it contains the trip from hub $h u b_i$ to itself.

Applications of the pHLRP may be found, for example, in public and urban transportation systems such as subway and train lines. For example, in subway and train transportation, the inter-hub route consists of platforms that are used to consolidate passengers from many origins and allow their delivery to many destinations. In addition, some of these platforms are served by bus lines which are actually local routes that connect the passengers to the platforms. Another application of this problem arises when designing distributed data networks.

Regarding solution approaches for the problem, there are three heuristic methods available so far [\[20](#page-18-8)]: (i) a multi-start variable neighborhood descent, (ii) a Biased Random-Key Genetic Algorithm (BRKGA), and (iii) Local Solver version 6.0. According to the documentation, Local Solver combines different search techniques, constraint propagation, and linear mixed-integer as well as nonlinear programming. The multi-start variable neighborhood descent, in each iteration, applies a variable neighborhood descent (VND) on a random initial solution. Two different variable neighborhood descents (VND) are proposed in [\[20\]](#page-18-8). They both explore the same neighborhood structures defined by inserting non-hub nodes from one tour to another; exchanging two non-hub nodes; and changing a hub node. In addition both use the Lin-Kernighan heuristic to improve tours. As search strategies, both VNDs use neither a first nor best improvement search strategy, but execute a random improvement move. However, they differ in the way the exploration of neighborhoods is organized. The first one uses a nested VND scheme, while the other performs consecutive nested Neighborhood Search (CNS). The multi-start VND based on nested VND is denoted by M-VND, while the one using CNS is denoted as M-CNS. Both M-VND and M-CNS exhibit better performance than BRKGA.

In this paper we propose a general variable neighborhood search (GVNS) based approach. GVNS also employs VND as a local search but does not use a completely random re-start to resolve local optimum traps as is the case in M-VND and M-CNS. Instead, it modifies the current solution to varying degrees in order to resolve a local optimum trap. In addition, our GVNS differs from M-VND and M-CNS in the VND procedure. The difference comes from the number of neighborhood structures, the way the neighborhoods are organized in VND, and the search strategy. More precisely, our VND explores seven neighborhood structures in a sequential way using the first improvement strategy. This is a completely different approach from that used in [\[20](#page-18-8)]. In addition, our VND does not use the Lin-Kernighan heuristic, but adopts instead other neighborhood structures applicable to the traveling salesman problem (TSP).

The proposed GVNS approach is validated on benchmark instances, and its performance demonstrates both a superior efficiency and effectiveness. Namely, our GVNS succeeds to establish a large number of new best-known solutions. In addition, it turns out to be very efficient in solving large problem instances in a short time. Namely, in almost all large instances it provides significantly better solutions than previous ones in a fraction of the time. Hence, the contributions of the paper may be summarized as:

- A new heuristic for pHLRP is proposed and tested. It is based on exploring seven neighborhoods within a GVNS framework.
- The proposed approach outperforms the previous ones from the literature to a great extent, particularly on large scale instances.
- 691 out of 912 best-known solutions are improved by the proposed GVNS heuristic;
- The algorithm is also very fast compared to the previous methods.

The rest of the paper is organized as follows. The next section details the different components of the new heuristic, and puts them all together within the GVNS framework. This is followed by extensive computational work, which includes a validation of the selected VND structure and parameter specifications, a comparison with the state-of-the-art, and a statistical analysis that confirms the superiority of the new heuristic. Lastly, we provide some conclusions and suggestions for future research.

2 Proposed variable neighborhood search approach

In this section, we provide a detailed description of our solution approach. The section starts by describing the solution representation and the procedure for generating an initial solution. After that, details are provided on the various neighborhood structures used to refine the solution. Finally, the main ingredients of the proposed general variable neighborhood search are described.

2.1 Solution representation

Any solution of the pHLRP may be represented by the three following components: set $H = \{hub_1, hub_2, \ldots, hub_n\}$ containing the p selected hub nodes, set $R =$ ${r_1, r_2, \ldots, r_p}$ containing the local routes assigned to the hubs, and the global route denoted as *hub*_*tour* which connects the hubs from set *H*. In the established notation each local route r_i originates at hub node $h u b_i$. In other words, each route r_i has the form $h u b_i - n_1^i - \cdots - n_k^i - h u b_i$ where $n_1^i - n_2^i \cdots - n_k^i$ is the ordered sequence of non-hub nodes visited along the route starting and ending at *hubi* . Hence, the solution is represented as the triplet $S = (H, hub\; tour, R)$.

2.2 Initial solution

To construct an initial solution we proceed in the following way. From the set *V* we choose p nodes at random to constitute set H . After that, we assign nodes to hubs and form routes as follows. For each non-hub node the corresponding hub is chosen at random from the set *H*, and the node is added at the end of the route of the chosen hub. The assignment is made while respecting the imposed constraint on the route length. The *hub_tour* is constructed as a tour that connects hubs $hub_1, hub_2,..., hub_p$ in this order. The outline of the procedure is given by Algorithm [1.](#page-4-0) Note that a feasible solution is generated in this way.

```
Algorithm 1: Generating an initial solution
     Function Initial_Solution(S);
   1 Choose randomly p nodes from V to constitute set
      H = \{hub_1, hub_2, \ldots, hub_n\};
   2 hub_tour ← hub<sub>1</sub> − hub<sub>2</sub> −··· − hub<sub>p</sub> − hub<sub>1</sub>; /*global tour*/;
   3 for i = 1 to p do
   4 \vert r_i \leftarrow hub_i - hub_i; /*creating local tours with just root hubs*/;
   5 length<sub>i</sub> = 1;
     end
   6 for each n ∈ V − H do
   7 Select at random i \in [1, p] such that length_i < C;<br>
8 r_i \leftarrow hub_i - \cdots - n - hub_i: /*adding node at the
        r_i \leftarrow hub_i - \cdots - n - hub_i; /*adding node at the end of current tour*/;
   9 \vert length<sub>i</sub> \leftarrow length<sub>i</sub> + 1;
     end
  10 return (H, hub_tour, R);
```
2.3 Neighborhood structures

The solution generated by procedure Algorithm [1](#page-4-0) will likely be far from the optimal one. However, it may be improved by replacing a hub by a node which is not a hub, or moving a non-hub node from one route to another, or changing the order of visiting

nodes in the local routes and in the global route. Accordingly, we define the following neighborhoods:

- *Insert neighborhood* (*Ninsert*)- This neighborhood is based on a move that takes a single non-hub node from one local route and inserts it in another local route, while respecting the capacity of the local routes.
- *Exchange* (*Nexhange*) This neighborhood is based on a move that exchanges two non-hub nodes from two different local routes. Note that these moves preserve the number of non-hub nodes in each local route.
- *Change hub* (*Nchange*_*hub*) This neighborhood is based on moves that replace a single hub node by a non-hub node from the local route originating at this hub.
- *LocalRoute neighborhoods* The purpose of LocalRoute neighborhoods is to optimize the order of visiting non-hub nodes within local routes as well as hub nodes within the global route. Here we adopt some k-opt neighborhoods used for solving the Traveling Salesman Problem (TSP) and its variants (see e.g. [\[12](#page-18-9)[,24](#page-18-10)[,28\]](#page-18-11)). They are: 1-opt, 2-opt and OR-opt-1. The 2-opt neighborhood is based on a move that inverts a part of the route. A simplified variant of the 2-opt move is the 1-opt move, which inverts a part of a tour consisting of 2 consecutive nodes. The ORopt-1 neighborhood is based on an insertion move that relocates a node from one position to another. Two types of insertion move may be distinguished: forward and backward insertion moves depending on whether a node is moved forward or backward in the tour, respectively. 1-opt, 2-opt, OR-opt-1 forward and OR-opt-1 backward will be denoted as N_{1opt} , N_{2opt} , N_{Oropt1} *for* and $N_{Oropt1back}$, respectively.

Note that all of these neighborhood allow feasible moves only, i.e., moves that respect the capacity constraints (i.e., do not violate the maximum length of a local route). The change in value of the objective function caused by performing any of these moves may be calculated in constant time O(1).

2.4 Variable Neighborhood descent

The above neighborhood structures do not necessarily yield the same local optima. That is, a local optimum with respect to one neighborhood structure is not necessarily a local optimum with respect to another. In order to efficiently exploit the defined neighborhood structures we embed them in a sequential variable neighborhood descent (VND) framework [\[7](#page-18-12)[,13](#page-18-13)[,24](#page-18-10)]. In this way, the VND takes a feasible solution at the input and as the output returns a solution which is a local optimum with respect to all employed neighborhood structures. The steps of the VND procedure are provided in Algorithm [2.](#page-6-0) As observed in previous works on VND (see e.g. $[24]$), the ordering of the neighborhood structures in the search as well as the search strategy may have a significant impact on the final performance of the VND procedure. Due to the large size of the neighborhood structures we selected a first improvement search strategy, i.e., as soon as an improving solution is detected in some neighborhood structure it replaces the incumbent solution, and the search is re-started from it. Regarding the order of neighborhoods within the VND, we then performed an empirical study to identify a most suitable one (see Sect. [3\)](#page-7-0). This empirical study provided the final selected order shown in line 1 of the algorithm. The statement $LocalSearch(S, N_k)$ in line 3 denotes the local search with respect to the kth neighborhood in the ordered list *N* of neighborhoods.

Algorithm 2: Basic Sequential Variable Neighborhood Descent

```
Function BVND(S);
    1 N = \{N_{insert}, N_{exchange}, N_{change\_hub}, N_{1opt}, N_{2opt}, N_{Oropt1back}, N_{Oropt1} \}2 k \leftarrow 1;
      while k \leq 7 do
    \begin{array}{c|c} 3 & S' \leftarrow \text{LocalSearch}(S, N_k); \\ 4 & k \leftarrow k+1. \end{array}k \leftarrow k + 1;
          if S	 better than S then
    5 \mid S \leftarrow S';\bullet | k \leftarrow 1;
          end
       end
7 return S
```
2.5 General variable neighborhood search

In this section we describe our *general variable neighborhood search* (GVNS) heuristic. Recall that GVNS is a variant of variable neighborhood search [\[13](#page-18-13)[,26\]](#page-18-14) that uses a *variable neighborhood descent* (VND) in the intensification phase (instead of a basic local search with a single neighborhood (see e.g., [\[2](#page-17-3)[,25\]](#page-18-15))) and the usual *shaking procedure* in the diversification phase. Our implementation employs the VND presented in Algorithm [2](#page-6-0) for the intensification (or improvement) phase, while at the diversification phase we employ the shaking procedure presented in Algorithm [3.](#page-6-1) The shaking procedure moves the search to new regions of the solution space by relocating non-hub nodes to different routes and changing the order of visitation of a single local route. The improvement phase and the shaking procedure together with the neighborhood change step are executed alternately until a predefined stopping criterion is met. Here the stopping criterion is given by a limit on CPU time (T_{max}) (see line 10 in Algorithm [4\)](#page-7-1). GVNS has an additional parameter k_{max} which denotes the maximum number of iterations (moves) that may be performed within a single shake (Algorithm [3\)](#page-6-1) and which affects the level of diversification obtained by the shaking procedure. The parameter k_{max} is determined empirically in our study (see Sect. [3\)](#page-7-0).

Algorithm 3: Shaking procedure

Function shaking(*S*,*k*); $\mathcal{N} = \{\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3\} = \{N_{insert}, N_{Oropt1}$ *for*, $N_{Oropt1back}\};$ **1 for** $i = 1$ **to** k **do 2** | Select *S*^{*''*} in $\mathcal{N}_1(S)$ at random; $3 \mid S \leftarrow S''$; **end ⁴ for** *i* = 1 *to k* **do 5** Select *S*^{*''*} in $\mathcal{N}_2(S) \cup \mathcal{N}_3(S)$ at random; $\mathbf{s} \quad S \leftarrow S''$; **end ⁷ return** *S*;

Algorithm 4: General VNS

```
Function GVNS(S, k_{max}, T_{max});
  1 S \leftarrow Initial Solution(S);
 2 repeat
 \begin{array}{c|c} 3 & k \leftarrow 1; \\ 4 & \text{while } k \end{array}4 while k \leq k_{max} do<br>5 do<br>S' \leftarrow shaking
  \begin{array}{c} 5 \ 6 \end{array} \begin{array}{c} S' \leftarrow \text{shaking}(S, k) \, ; \\ S'' \leftarrow \text{BVID}(S') \, ; \end{array}\bullet | | S'' \leftarrow BVND (S');
  \begin{array}{|c|c|c|c|}\n\hline\n\text{A} & k & k+1; \\
\hline\n\text{B} & \text{B} & \text{B} & \text{B} \\
\hline\n\end{array}if S'' is better then S then
  9 \vert \vert S \leftarrow S''; k \leftarrow 1;
                 end
           end
10 \mid T \leftarrow CpuTime();
     until T > T_{max};
11 Return S;
```
3 Computational results

In this section, we present the results obtained by our GVNS heuristic on benchmark instances from the literature. In addition, we present the results of preliminary testing used to determine the order of neighborhoods within the VND, as well as to determine an appropriate value of the parameter k_{max} . Also note that a CPU time limit T_{max} of 60 s is imposed on all runs of the GVNS heuristic.

The experiments are conducted on a machine with an Intel (R) Core (TM) i5-3470 CPU @ 3.20 GHz and 16 GB of RAM.

3.1 Benchmark instances

For testing purposes, the benchmark instances originally proposed in [\[20\]](#page-18-8) are used. The benchmark set is derived from 76 TSPLIB instances by varying the value of parameter α and considering different scenarios on the tour length C, and the requested number of hubs *p*. More precisely, the value of the parameter α is taken from the set {0.2, 0.4, 0.6, 0.8}, while three scenarios are considered for values of *C* and *p*:

- *Scenario ST p* = $[0.2|V|]$ and $C = \lceil \frac{|V|}{p} \rceil$. In this case the local routes are forced to be non-empty and tight.
- *Scenario SL* $p = [0.2|V|]$ and $C = \lfloor 1.8 \lceil \frac{|V|}{p} \rceil \rfloor$. Here some local routes may be very large and some may be empty.
- *Scenario SQ* $p = C = \sqrt{|V|}$ *.* Here local routes may be large or empty, but not as loose as scenario SL .

Hence, in total we have $76 \times 4 \times 3 = 912$ test instances. The instances with 100 nodes or less are classified as small, while the remaining ones are large.

3.2 Determining the best order of neighborhoods

The aim of this section is to determine the best order of neighborhoods inside our VND procedure. Due to the large number of neighborhood structures considered here, examining all possible orders would be far too time consuming. For this reason, we divide the neighborhood structures into two groups: hub-location neighborhoods and LocalRoute neighborhoods. Hub-location neighborhoods are those that concern the selection of hubs and node-hub assignments, while LocalRoute neighborhoods are those that "optimize" local tours. Having such a division of neighborhoods, we are now interested to determine the best order of neighborhood groups. The order of neighborhoods within the set of hub-location neighborhoods may be determined so that those based on changing node-hub allocations are examined first followed by those that change the hubs. Such a way of exploring neighborhoods is suggested for example in [\[16](#page-18-16)]. Hence, respecting the rule that the neighborhoods of smaller cardinality are examined first, we obtain the following order of neighborhoods within the hub-location group: $N_{hub} := \{N_{insert}, N_{exchange}, N_{change_hub}\}.$ In similar fashion, we order neighborhood structures within the LocalRoute neighborhood group as $N_{Local Route} := \{N_{1opt}, N_{2opt}, N_{Oropt1back}, N_{Oropt1} \}$. To validate this order we note the following observations: the lowest cardinality neighborhood is examined first; the 2-opt neighborhood should be examined before the OR-opt neighborhoods (e.g., as noted in [\[24](#page-18-10)]); ORopt_backward should be examined before ORopt_forward (see e.g., $[12, 27, 32]$ $[12, 27, 32]$).

Now that the order of neighborhoods within each group is specified, we examine the best order of groups within the VND paradigm. Hence, we consider the following orders: 1) hub-location neighborhoods, LocalRoute neighborhoods; 2) LocalRoute neighborhoods; hub-location neighborhoods, and 3) hybrid order, where after each neighborhood in the hub-location group we apply the LocalRoute neighborhoods. To differentiate these three VNDs we designate them as:

- VND_1-order of neighborhoods: *NLocal Route*, *Nhub*;
- VND_2-order of neighborhoods: *Nhub*, *NLocal Route*;
- VND_3-order of neighborhoods: *Ninsert*, *NLocal Route*, *Nexchange*, *NLocal Route*, *Nchange*_*hub*, *NLocal Route*;

To assess the performance of the above three VNDs, each one is executed 50000 times on a selected subset of instances, each time starting from a different initial solution generated at random as described in Algorithm [1.](#page-4-0) For testing purposes, we selected a subset of 240 instances with different topologies. This subset is based on the following 20 TSPlib instances: a280, li535, att532, bier127, brg180, ch130, d198, d493, fl417, gil262, gr137, gr202, gr229, gr431, kroA150, kroA200, lin318, pa561, pcb442, pr144. Each of these instances is considered under 3 different scenarios and 4 different values of the parameter α . For each VND, the average results taken over 20 instances, all using the same scenario and the same value of parameter α , are provided in Table [1.](#page-10-0) More precisely, we report the averages of the best, the average and the worst solution values (Columns 'Best', 'Avg.' and 'Worst', respectively), as well as the average time-to-target (Columns 'time') in seconds. The time-to-target represents

the CPU time consumed by a method to reach a final solution value reported at the output.

From the reported results we observe, that average times-to-target of all three VNDs are almost the same. Regarding solution quality, VND_1 exhibits the poorest performance. Comparing VND_2 and VND_3, we observe that in all cases but one (see the line corresponding to the scenario SL and $\alpha = 0.8$), VND_2 provides better (best) solutions. In addition, in 9 out of 12 cases, we see that the averages of best, average and worst solution values provided by VND_2 are better than the corresponding ones of VND_3. Consequently, VND_2 may be identified as the best VND. This further implies that the best order of neighborhoods is:

Ninsert, *Nexchange*, *Nchange*_*hub*, *N*1*opt*, *N*2*opt*, *NOropt*1*back* , *NOropt*¹ *f or*.

3.3 Assessing the value of parameter *kmax*

After detecting the best order, we needed to also tune parameter k_{max} , which determines the level of diversification that will be used within GVNS to resolve local optimum traps. It is known that for larger values of parameter k_{max} , the corresponding VNS heuristic tends to behave as a random multi-start heuristic (i.e., a heuristic that generates a completely random solution at each re-start that does not retain any features of the current best solution). In order to avoid this worst-case scenario, we want to keep the k_{max} value as small as possible while having at the same time high quality results. Since the number of local routes in a solution is expressed by the parameter *p*, we try to keep the number of local routes affected by the shaking procedure to be as small as possible by attempting the following settings: $k_{max} = min(5, p)$, k_{max} = min(10, *p*) and k_{max} = min(20, *p*). In all testing, the time limit parameter T_{max} is set to 60 s as noted above, and for each choice of k_{max} GVNS is executed 20 times, each time starting from a different initial solution generated by Algorithm [1.](#page-4-0) For testing purposes, we used the same subset of 240 instances as described above. The average results taken over 20 instances, all using the same scenario and the same value of parameter α , are provided in Table [2.](#page-11-0) The column headings are defined in the same way as previously described.

From the reported results, it follows that for scenarios ST and SL the best choice is $k_{max} = min(10, p)$, regardless of the value of α . This choice of the value of parameter *kmax* , yields better best and average solution values than the other choices. On the other hand for the SQ scenario the best choice turns out to be $k_{max} = \min(5, p)$. Hence, in all subsequent testing, we will use $k_{max} = min(10, p)$ for the ST and SL scenarios, and $k_{max} = min(5, p)$ for the scenario SQ. It should also be emphasized that regardless of the value of k_{max} , our GVNS exhibits a good stability with respect to time-to-target.

An interesting observation may be deduced after comparing the results reported in Table [1](#page-10-0) (multi-start VND) and Table [2](#page-11-0) (GVNS). Namely, in all test cases, the solutions obtained by the proposed GVNS are much better than the ones reached by the multi-start VND. Hence, it follows that the proposed shaking procedure resolves local optimum traps more effectively than the random re-start used within the multistart VND.

3.4 Comparison with the state-of-the-art

In this section we compare our GVNS against the state-of-the-art heuristics presented in [\[20](#page-18-8)]: M-VND, M-CNS, and LocalSolver. In all tests the parameters of our GVNS are set as previously described, i.e., $T_{max} = 60$ s, $k_{max} = min(10, p)$ on scenarios ST and SL, and $k_{max} = min(5, p)$ on scenario SQ. As an improvement routine, VND₋₂ is used. On each of the 912 instances tested GVNS is executed 20 times, each time starting from a different initial solution generated by Algorithm [1.](#page-4-0)

The detailed comparison may be found at the following webpage: [http://www.mi.](http://www.mi.sanu.ac.rs/~nenad/p-hubLR/) [sanu.ac.rs/~nenad/p-hubLR/,](http://www.mi.sanu.ac.rs/~nenad/p-hubLR/) while here we present the summary results. All instances of the same size type (small or large), tested under the same scenario (ST, SL or SQ) and the same value of the parameter $\alpha \in \{0.2, 0.4, 0.6, 0.8\}$ are considered as a test case. Hence, for each test case, we report the average results over all instances in that test case. Table [3](#page-13-0) reports the results on small instances, while Table [4](#page-14-0) provides results on large instances. The column headings are defined as follows. Columns 'Scenario' and ' α ' provide information on the scenario considered and the value of parameter α , respectively. The next column 'Best Known' provides the average of best-known solution values obtained from the literature for each test case. After that, in the next columns, for each of methods M-VND, M-CNS, and LocalSolver (LS), we report the average percentage deviations of their solution values from the best-known solution values and the average time-to-target. For our GVNS, we report the averages of the best and average solution values on each test case, the average time-to-target and the average percentage deviations of the best found solution values from the current best known values. For each method, the percentage deviation is calculated as:

$$
\frac{value - best_known}{best_known} \times 100,
$$

where v*alue* stands for the solution value found by the considered method and *best*_*kno*w*n* is the best-known solution value reported in the literature [\[20\]](#page-18-8). Note that the percentage deviations for M-VND, M-CNS, and LocalSolver are taken directly from [\[20\]](#page-18-8) as well as all other data related to these methods.

Table [5](#page-15-0) provides the following information for each test case: number of instances where the proposed GVNS establishes a new best-known solution (columns 'Better'); number of instances where it reaches the previous best-known solution (columns 'Same'); and number of instances where it fails to improve or reach the previous best-known solution (columns 'Worse').

From the results reported on small size instances in Table [3,](#page-13-0) we observe that our GVNS provides very good solution values in short time. On all 12 test cases, the averages of best and average solution values are less than the averages of best-known values. The previous best-known solution values are improved up to 0.73%. Just on two test cases (SL, 0.20 and SL, 0.80), our GVNS provides higher percentage deviation from the best known solution values than LS. In the remaining 10 cases our GVNS not only provides less percentage deviation than LS but also improves previous bestknown solution values. The total number of new best-known solutions established by our GVNS is 135 (see Table [5\)](#page-15-0). On 138 instances, our GVNS matches the previous

 ces

 \mathcal{D} Springer

Scenario	α	Small instances			Large instances		
		Better	Same	Worse	Better	Same	Worse
ST	0.2	11	16	1	47	$\mathbf{0}$	
	0.4	9	19	$\overline{0}$	47	Ω	
	0.6	10	17	1	47	Ω	1
	0.8	12	16	θ	46	Ω	\mathfrak{D}
SL	0.2	12	$\mathbf{0}$	16	47	Ω	
	0.4	13	1	14	47	Ω	
	0.6	12	3	13	43	Ω	5
	0.8	12	$\overline{2}$	14	47	Ω	
SQ	0.2	11	14	3	45	Ω	3
	0.4	13	15	$\mathbf{0}$	46	Ω	\overline{c}
	0.6	10	17	1	47	Ω	
	0.8	10	18	θ	47	Ω	
Total		135	138	63	556	$\overline{0}$	20

Table 5 Number of instances where GVNS provides a better, same or worse solution than best-known from the literature

best-known solutions, while on 63 instances it fails to improve or match the previous best-known solutions.

Comparing the results on large scale instances, it follows that our GVNS exhibits better performance than any other solution approach. On every instance it either improves or matches the previous best-known solution. The average percentage improvement is in the range of 5.32 and 11%. It should be emphasized that our GVNS consumes a lot less CPU time-to-target than any other method. M-VND has very high percentage deviation from the previous best-known solution values, reaching as high as 14.51%. M-CNS and LS exhibit better performance than M-VND, but still worse than GVNS. For example, on test case SL 0.60, where percentage deviations of M-CNS and LS are both 4.51%, our GVNS improves best-known solutions by 6.52%. Moreover, our GVNS improves the previous best-known solutions on 556 out of 576 test instances (see Table [5\)](#page-15-0). It should be emphasized that, as reported in [\[20](#page-18-8)], M-CNS and LS exhibit the same performance on all large instances.

Overall, we see that GVNS improves best known solution values on 691 out of 912 instances; matches the previous best-known solution values on 138 instances, while on only 83 instances it neither improves nor reaches the previous best-known solution values. Thus as of now, 829 out of 912 best-known solutions are attributed to the proposed GVNS, i.e., around 90% of the current best-known solutions. In addition, it follows that the proposed GVNS is highly suitable for solving large scale instances in a short time by providing high quality solutions within 60 s. This further implies that our GVNS successfully responds to the main requirement of a heuristic approach, that is, to provide high quality solutions in short time.

Scenario	α		Small instances			Large instances		
		M-VND	M-CNS	LS	M-VND	M-CNS	LS	
ST	0.2	5.35E-04	5.35E-04	1.22E-02	1.63E-09	3.06E-09	3.06E-09	
	0.4	1.22E-04	2.93E-04	3.90E-03	1.63E-09	5.03E-09	5.03E-09	
	0.6	6.10E-05	2.92E-04	2.00E-03	1.63E-09	4.18E-09	4.18E-09	
	0.8	2.93E-04	6.10E-05	4.88E-04	2.40E-09	7.56E-09	7.56E-09	
SL.	0.2	3.97E-02	2.50E-03	9.23E-01	1.63E-09	2.56E-08	2.56E-08	
	0.4	3.06E-02	7.40E-03	7.19E-01	1.63E-09	1.41E-08	1.41E-08	
	0.6	9.40E-03	1.28E-02	7.37E-01	1.97E-09	4.05E-07	4.05E-07	
	0.8	7.40E-03	1.99E-02	5.01E-01	1.63E-09	9.26E-09	9.26E-09	
SQ	0.2	1.60E-03	1.00E-03	2.45E-02	1.63E-09	4.08E-08	4.08E-08	
	0.4	1.22E-04	1.32E-04	2.44E-04	1.63E-09	1.41E-08	1.41E-08	
	0.6	3.05E-04	3.05E-04	3.05E-04	1.63E-09	8.72E-09	8.72E-09	
	0.8	4.38E-04	1.22E-04	2.00E-03	1.63E-09	1.25E-08	1.25E-08	

Table 6 *p*-values of the Wilcoxon signed-rank test

3.5 Statistical tests

The aim of this section is to determine if the difference in performance between our GVNS and each other method in the comparison is statistically significant or not. In this regard, the Wilcoxon signed-rank test is performed with a significance level of 0.05. The outcome of the tests is provided in Table [6.](#page-16-0) For each test case, the corresponding row provides *p*-values obtained after applying the Wilcoxon signed-rank test on the results found by the proposed GVNS and the methods shown in the column headings. A *p*-value less than 0.05 indicates that there is a significant difference between the two methods, otherwise there is insufficient evidence. From the table we see that on large instances all *p*-values are orders of magnitude less than 0.05, which implies that there is a significant difference between the proposed GVNS and each of the previously proposed methods. Since, on large instances, the proposed GVNS provides better solution values than any existing method (see Table [4\)](#page-14-0), it may be inferred that the GVNS is the superior method on large size instances, in terms of the solution quality. Similar conclusions may be drawn when comparing GVNS with M-VND and M-CNS on small instances. When comparing the GVNS and LS on small instances, the previous conclusions are applicable on all test cases except for small instances and scenario SL, where no significant differences between GVNS and LS are observed.

4 Conclusions

In this paper an established hub location routing problem is studied, and an efficient solution approach is proposed. The new solution approach follows the general variable neighborhood search framework where in our case seven different neighborhood structures are identified to exploit the nature of the problem. These seven neighborhoods are embedded in a sequential variable neighborhood descent scheme in an efficient and effective ordering determined from a detailed empirical study. In addition, all other parameters and components of our approach are carefully tailored with the aim to obtain high-quality solutions in a short time. Comparing our heuristic with the stateof-the-art approaches reveals that we succeeded to do so to a great extent. Our heuristic established 691 new best-known solutions on a benchmark set of 912 instances while consuming orders of magnitude less CPU time than all competitors. In addition, 829 out of 912 current best-known solutions can be attributed to the proposed GVNS. The statistical tests we conducted strongly confirm the superiority of our approach over the previous ones.

The SL small-sized instances with poor GVNS performance should be the subject of future work in order to detect possible anomalies in the solution space of these instances. As additional future research directions, we suggest the hybridization of the proposed solution method with some exact approaches to hopefully obtain guaranteed optimal solutions, as well as extending the study to other hub location routing problems. Some variants to be studied include for example, capacitated variants of the current problem, with capacities on the edges or at the hubs, allowing the assignment of non-hub nodes to more than one hub, and allowing for different configurations such as a tree structure for inter-hub connections.

In addition, some recent works use an adaptive search strategy, which combines first and best improvement strategies. More specifically, the best improvement is selected for small- and medium-sized instances, and the first improvement for the solution of large problem instances [\[18](#page-18-18)[,19\]](#page-18-19). Therefore, this approach may be also a promising future work direction. Moreover, in [\[19\]](#page-18-19) new adaptive shaking approaches can be found, which may further enhance the efficiency of GVNS. Hence, this is one more possible research direction. Another possible research direction may be to examine the use of an adaptive neighborhoods' ordering mechanism of neighborhood structures as was done in [\[33](#page-19-3)], for example.

Acknowledgements This research is supported by ELSAT2020 (Eco-mobilité, Logistique, Scurité and Adaptabilité des Transports l'horizon 2020) Project co-financed by European Union, France and Région Hauts-de-France.

References

- 1. Alumur, S., Kara, B.Y.: Network hub location problems: the state of the art. Eur. J. Oper. Res. **190**(1), 1–21 (2008)
- 2. Amaldass, N.I.L., Lucas, C., Mladenovic, N.: Variable neighbourhood search for financial derivative problem. Yugosl. J. Oper. Res. **29**(3), 359–373 (2019)
- 3. Brimberg, J., Mišković, S., Todosijević, R., Urošević, D.: The uncapacitated r-allocation p-hub center problem. Int. Trans. Oper. Res. (2020). <https://doi.org/10.1111/itor.12801>
- 4. Brimberg, J., Mladenović, N., Todosijević, R., Urošević, D.: A basic variable neighborhood search heuristic for the uncapacitated multiple allocation p-hub center problem. Optim. Lett. **11**(2), 313–327 (2017)
- 5. Brimberg, J., Mladenović, N., Todosijević, R., Urošević, D.: General variable neighborhood search for the uncapacitated single allocation p-hub center problem. Optim. Lett. **11**(2), 377–388 (2017)
- 6. Brimberg, J., Mladenović, N., Todosijević, R., Urošević, D.: A non-triangular hub location problem. Optim. Lett. (2019). <https://doi.org/10.1007/s11590-019-01392-2>
- 7. Duarte, A., Sánchez-Oro, J., Mladenović, N., Todosijević, R.: Variable neighborhood descent. In: Martí, R., Pardalos, P.M., Resende, M.G.C. (eds.) Handbook of Heuristics, pp. 341–367. Springer, Cham (2018)
- 8. Ernst, A.T., Hamacher, H., Jiang, H., Krishnamoorthy, M., Woeginger, G.: Uncapacitated single and multiple allocation p-hub center problems. Comput. Oper. Res. **36**(7), 2230–2241 (2009)
- 9. Ernst, A.T., Jiang, H., Krishnamoorthy, M., Baatar, H.: Reformulations and computational results for uncapacitated single and multiple allocation hub covering problems. Work. Pap. Ser. **1**, 1–18 (2011)
- 10. Farahani, R.Z., Hekmatfar, M., Arabani, A.B., Nikbakhsh, E.: Hub location problems: a review of models, classification, solution techniques, and applications. Comput. Ind. Eng. **64**(4), 1096–1109 (2013)
- 11. Gavriliouk, E.O.: Aggregation in hub location problems. Comput. Oper. Res. **36**(12), 3136–3142 (2009)
- 12. Gelareh, S., Gendron, B., Hanafi, S., Monemi, R.N., Todosijević, R.: The selective traveling salesman problem with draft limits. J. Heuristics **26**, 339–352 (2020)
- 13. Hansen, P., Mladenović, N., Todosijević, R., Hanafi, S.: Variable neighborhood search: basics and variants. EURO J. Comput. Optim. **5**(3), 423–454 (2017)
- 14. Hoff, A., Peiró, J., Corberán, Á., Martí, R.: Heuristics for the capacitated modular hub location problem. Comput. Oper. Res. **86**, 94–109 (2017)
- 15. Hwang, Y.H., Lee, Y.H.: Uncapacitated single allocation p-hub maximal covering problem. Comput. Ind. Eng. **63**(2), 382–389 (2012)
- 16. Ilić, A., Urošević, D., Brimberg, J., Mladenović, N.: A general variable neighborhood search for solving the uncapacitated single allocation p-hub median problem. Eur. J. Oper. Res. **206**(2), 289–300 (2010)
- 17. Janković, O., Mišković, S., Stanimirović, Z., Todosijević, R.: Novel formulations and vns-based heuristics for single and multiple allocation p-hub maximal covering problems. Ann. Oper. Res. **259**(1–2), 191–216 (2017)
- 18. Karakostas, P., Sifaleras, A., Georgiadis, M.C.: A general variable neighborhood search-based solution approach for the location-inventory-routing problem with distribution outsourcing. Comput. Chem. Eng. **126**, 263–279 (2019)
- 19. Karakostas, P., Sifaleras, A., Georgiadis, M.C.: Adaptive variable neighborhood search solution methods for the fleet size and mix pollution location-inventory-routing problem. Expert Syst. Appl. **153**, 113444 (2020)
- 20. Lopes, M.C., de Andrade, C.E., de Queiroz, T.A., Resende, M.G., Miyazawa, F.K.: Heuristics for a hub location-routing problem. Networks **68**(1), 54–90 (2016)
- 21. Martí, R., Corberán, Á., Peiró, J.: Scatter search for an uncapacitated p-hub median problem. Comput. Oper. Res. **58**, 53–66 (2015)
- 22. Meyer, T., Ernst, A.T., Krishnamoorthy, M.: A 2-phase algorithm for solving the single allocation p-hub center problem. Comput. Oper. Res. **36**(12), 3143–3151 (2009)
- 23. Mikić, M., Todosijević, R., Urošević, D.: Less is more: general variable neighborhood search for the capacitated modular hub location problem. Comput. Oper. Res. **110**, 101–115 (2019)
- 24. Mjirda, A., Todosijević, R., Hanafi, S., Hansen, P., Mladenović, N.: Sequential variable neighborhood descent variants: an empirical study on the traveling salesman problem. Int. Trans. Oper. Res. **24**(3), 615–633 (2017)
- 25. Mladenović, N., Alkandari, A., Pei, J., Todosijević, R., Pardalos, P.M.: Less is more approach: basic variable neighborhood search for the obnoxious p-median problem. Int. Trans. Oper. Res. **27**(1), 480– 493 (2020)
- 26. Mladenović, N., Hansen, P.: Variable neighborhood search. Comput. Oper. Res. 24(11), 1097-1100 (1997)
- 27. Mladenović, N., Todosijević, R., Urošević, D.: An efficient gvns for solving traveling salesman problem with time windows. Electron. Notes Discrete Math. **39**, 83–90 (2012)
- 28. Mladenović, N., Todosijević, R., Urošević, D.: An efficient general variable neighborhood search for large travelling salesman problem with time windows. Yugosl. J. Oper. Res. **23**(1), 19–30 (2013)
- 29. Peiró, J., Corberán, A., Martí, R.: Grasp for the uncapacitated r-allocation p-hub median problem. Comput. Oper. Res. **43**(1), 50–60 (2014)
- 30. Peker,M., Kara, B.Y.: The p-hub maximal covering problem and extensions for gradual decay functions. Omega **54**, 158–172 (2015)
- 31. Talbi, E.G., Todosijević, R.: The robust uncapacitated multiple allocation p-hub median problem. Comput. Ind. Eng. **110**, 322–332 (2017)
- 32. Todosijević, R., Mjirda, A., Mladenović, M., Hanafi, S., Gendron, B.: A general variable neighborhood search variants for the travelling salesman problem with draft limits. Optim. Lett. **11**(6), 1047–1056 (2017)
- 33. Todosijević, R., Mladenović, M., Hanafi, S., Mladenović, N., Crévits, I.: Adaptive general variable neighborhood search heuristics for solving the unit commitment problem. Int. J. Electr. Power Energy Syst. **78**, 873–883 (2016)
- 34. Todosijević, R., Urošević, D., Mladenović, N., Hanafi, S.: A geneal vaiable neighbohood seach fo solving the uncapacitated r-allocation p-hub median poblem. Optim. Lett. **11**(6), 1109–1121 (2017)
- 35. Weng, K., Yang, C., Ma, Y.: Two artificial intelligence heuristics in solving multiple allocation hub maximal covering problem. In: Huang, D.S., Li, K., Irwin, G.W. (eds.) Intelligent Computing, pp. 737–744. Springer, Berlin, Heidelberg (2006)
- 36. Yaman, H.: Allocation strategies in hub networks. Eur. J. Oper. Res. **211**(3), 442–451 (2011)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Mustapha Ratli¹ · Dragan Urošević² · Abdessamad Ait El Cadi¹ · Jack Brimberg³ · Nenad Mladenović⁴ · Raca Todosijević^{[1](http://orcid.org/0000-0002-9321-3464)} ®

E Raca Todosijević raca.todosijevic@uphf.fr

> Mustapha Ratli mustapha.ratli@uphf.fr

> Dragan Urošević draganu@mi.sanu.ac.rs

Abdessamad Ait El Cadi abdessamad.aitelcadi@uphf.fr

Jack Brimberg Jack.Brimberg@rmc.ca

Nenad Mladenović nenadmladenovic12@gmail.com

- ¹ LAMIH UMR CNRS 8201-Université Polytechnique Hauts-de-France, 59313 Valenciennes Cedex 9, France
- ² Mathematical Institute, Serbian Academy of Sciences and Arts, Belgrade, Serbia
- ³ Royal Military College of Canada, Kingston, ON, Canada
- ⁴ Khalifa University, Abu Dhabi, UAE