J Ind Eng Int (2014) 10:185–197
DOI 10.1007/s40092-014-0076-4

ORIGINAL RESEARCH

A modified NSGA-II solution for a new multi-objective hub maximal covering problem under uncertain shipments

Amir Ebrahimi Zade · Ahmad Sadegheih · Mohammad Mehdi Lotfi

Received: 12 October 2013 / Accepted: 30 June 2014 / Published online: 30 July 2014

Abstract Hubs are centers for collection, rearrangement, and redistribution of commodities in transportation networks. In this paper, non-linear multi-objective formulations for single and multiple allocation hub maximal covering problems as well as the linearized versions are proposed. The formulations substantially mitigate complexity of the existing models due to the fewer number of constraints and variables. Also, uncertain shipments are studied in the context of hub maximal covering problems. In many real-world applications, any link on the path from origin to destination may fail to work due to disruption. Therefore, in the proposed bi-objective model, maximizing safety of the weakest path in the network is considered as the second objective together with the traditional maximum coverage goal. Furthermore, to solve the bi-objective model, a modified version of NSGA-II with a new dynamic immigration operator is developed in which the accurate number of immigrants depends on the results of the other two common NSGA-II operators, i.e. mutation and crossover. Besides validating proposed models, computational results confirm a better performance of modified NSGA-II versus traditional one.

Keywords Facility location · Mathematical modeling · Hub maximal covering · Uncertainty · NSGA-II · Immigration operator

A. Ebrahimi Zade · A. Sadegheih · M. M. Lotfi (✉)
Department of Industrial Engineering, Yazd University, Saffayieh, Yazd, Iran
e-mail: Lotfi@yazd.ac.ir

A. Ebrahimi Zade
e-mail: aebrahimyzade@yahoo.com

A. Sadegheih
e-mail: sadegheih@yazd.ac.ir

Introduction

There are three major decisions in a supply chain design: location, routing and inventory decisions (Tavakkoli-Moghaddam et al. 2013). Hub location is one of the most appealing fields in facility location which attracted many researchers in recent years. Hub location problems have many applications in areas such as postal delivery systems, telecommunication networks, airline networks, and other delivery systems with numerous demand and supply nodes. Hubs are facilities which collect the flows from several origin centers, then rearrange and distribute them to their destinations. The effective use of hub nodes decreases the number of required links for connecting origins to destinations and helps to benefit the economies of scale. Hub location problem was first introduced by O'Kelly (1986). Afterwards, O'Kelly (1987) proposed the first mathematical model for hub location problem. Hub location problem consists of subcategories such as hub median, hub center, and hub covering problems (Alumur et al. 2012). This paper focuses on hub covering problems. The interested reader is advised to review the papers by Campbell and O'Kelly (2012) and Farahani et al. (2013) to study the other subcategories.

Hub covering problems, as location-allocation problems, consist of two sub problems namely hub set covering problem (HSCP) and hub maximal covering problem (HMCP). While HSCP is aimed at minimizing the transportation and hub establishment costs without any limitation on the number of established hubs, a hub maximal covering problem is constrained by the number of established hubs as an exogenous parameter. The hubs should be located in such a way that the total utility gained from all the covered origin/destination (O/D) pairs is maximized. Campbell (1994) introduced hub covering problems and proposed mathematical models for both HSCP and HMCP.

Kara and Tansel (2003) developed a non-linear model for the single allocation HSCP. They proved NP hardness of the problem and linearized it in three different ways. Wagner (2008) formulated a new mathematical model for HSCP, in which the cost discount factor is independent from the number of transmitted commodities. Tan and Kara (2007) used hub covering problem in the Turkish cargo delivery system including 81 cities with expert-based weights. Calık et al. (2009) studied a single allocation hub covering problem with an incomplete hub network. They proposed an efficient heuristic algorithm based on Tabu Search. Qu and Weng (2009) employed a path re-linking approach to solve HMCP. Karimi and Bashiri (2011) proposed models for HSCP and HMCP under a different coverage type and developed two heuristic algorithms to solve them. Fazel Zarandi et al. (2012) investigated a multiple allocation HSCP. What they assumed is that a node will be covered if there are at least a given number of paths to satisfy its demand. Furthermore, to consider the dispersion among hubs, they forced a lower bound on the distance between the established hubs. Hwang and Lee (2012) proposed two heuristics for HMCP and implemented them on the CAB dataset. The results confirmed satisfying performance of the heuristics in terms of both the solution quality and the computational time. In this paper, to mitigate the complexity of models for single and multiple allocation HMCPs, new mathematical formulations with fewer constraints and variables than the existing ones are proposed.

As mentioned earlier, most of the existing hub location models have been formulated in deterministic environments leading to invalid results for implementation; because, in practice, many problem parameters are characterized with high uncertainty. To face such uncertainty, some efforts have been made. Some of the major extensions to hub location problems under uncertainty are summarized in Table 1. As observed, demand, transportation time, customer's entrance rate to hubs, fixed costs of establishing hubs, covering radius and location of demand nodes were assumed to be uncertain in previous researches. Notably, in previous researches, uncertainty in the links among node pairs due to disruption has not been investigated. So, we study the effect of disruption in the links transmitting loads among node pairs in a hub network. Sometimes, to model the hub networks, it is critical to take into account link failure probability. For example, in martial distribution systems, there is usually a disruption probability for the transmitted cargos in war. According to some factors like airplane specifications as well as geographical and weather conditions of the path in air transportation systems, there are always some risks for flight. Also, when a massage is transmitted across different stations in a telecommunication system, it may be altered or destructed due to an interaction or crosstalk. Therefore, we develop a bi-objective HMCP model in which besides the coverage, safety of the weakest path in the obtained network is also maximized. Noteworthy, the safety of shipment through each link follows, as usual, a Bernoulli distribution independently from the other ones.

Table 1 Major contributions to hub location problems under uncertainty

References	Uncertain parameter	Solution approach
Sim et al. (2009)	Transportation time	Heuristic approach
Yang (2009)	Demand	Mixed-integer programming
Contreras et al. (2011)	Transportation time (dependent) Demand Transportation time (independent)	Monte Carlo simulation/ Benders decomposition
Mohammadi et al. (2011)	Entrance rate to hubs	Imperialist competitive algorithm (ICA)
Zhai et al. (2012)	Demand	Stochastic optimization
Alumur et al. (2012)	Fixed establishment costs demand	Robust optimization stochastic optimization
Eydi and Mirakhorli (2012)	Transportation time covering radius	Fuzzy linear programming
Mohammadi et al. (2013)	Transportation time	Multi-objective imperialist competitive algorithm
Davari et al. (2013)	Location of demand nodes	Fuzzy simulation

Based on the above explanations, main contributions of this paper are as follows:

1. Proposing new efficient mathematical formulations for single and multiple allocation HMCPs.
2. Investigating HMCPs under uncertain shipments by developing a bi-objective model maximizing safety of the paths in designed network.
3. Modifying the traditional NSGA-II to obtain a better performance in solving the HMCP problems.

The rest of the paper is organized as follows. In Sect. 2, we present the proposed mathematical formulations. To solve the proposed model efficiently, a modified version of NSGA-II is proposed in Sect. 3. Computational results to validate the models and analyze performance of the proposed algorithm are provided in Sect. 4. Finally, concluding remarks and some guidelines for further research are provided in Sect. 5.

Proposed mathematical model

In this section, mathematical formulations are developed for single and multiple allocation HMCPs as well as bi-objective HMCPs under uncertain shipments. Consider $N = \{1, 2, \ldots, n\}$ as the set of nodes, i and j as indices for origins (O) and destinations (D), respectively, and k and l as indices for hubs. Each node is a potential candidate for establishing a hub. Each O/D pair can be connected only through hubs. It is assumed that hub network is complete, and an O/D pair may be connected through one or two hubs. Hence, if at least one of the origin or destination nodes is a hub, it is possible to connect them directly. Moreover, it is assumed that traveling times among the nodes are symmetric and follow the triangle inequality (e.g., for i, k, j: $t_{ij} \le t_{ik} + t_{kj}$).

We assume that safety of the transmitted load between nodes i and j, independent of the other links, follows a Bernoulli distribution with parameter p_{ij}. Therefore, if it is planned to transmit load from origin i to hub k ($x_{ik} = 1$), hub k to hub l ($x_{kl} = 1$) and hub l to destination j ($x_{jl} = 1$), safety for transmitting the load between origin i and destination j will be equal to $p_{ik}.p_{kl}.p_{lj}$.

As special facilities are employed for transportation among the hubs, a cost discount factor α is introduced. Campbell (1994) suggested that O/D pair i–j will be covered through hubs k and l in the following three ways:

1. Total transportation cost (time or distance) from origin i to destination j via hubs k and l is less than a predetermined amount T ($c_{ik} + c_{kl} + c_{lj} \le T$). In this paper, this rule is used for covering an O/D pair.
2. Transportation cost (time or distance) for each link in the path from i to j via hubs k and l is less than a predetermined amount θ ($c_{ik} \le \theta, c_{kl} \le \theta, c_{lj} \le \theta$).
3. Transportation cost (time or distance) from origin i to hub k and hub l to destination j is less than a predetermined amount Δ ($c_{ik} \le \Delta, c_{lj} \le \Delta$).

Model parameters

- w_{ij} Importance of covering O/D pair i–j
- C_{ik} Traveling time (cost or distance) from node i to k
- T Maximum permissible transportation cost (time or distance) for covering O/D pairs
- P Number of hubs to be established
- p_{ik} Safety for the link transmitting loads between nodes i and k
- M A big number

Model variables

- x_{ik} Binary variable which is equal to 1 if node i is connected to hub k
- y_{ij} Binary variable which is equal to 1 if O/D pair i, j are connected to the hub network
- s Safety of the weakest path in the designed hub network

Single allocation HMCP model

In the following, the proposed model is presented for single allocation HMCPs:

$$\max \sum_i \sum_j \sum_k \sum_l w_{ij} x_{ik} x_{lj} \quad (1)$$

$$x_{ik} \le x_{kk};\ \forall i, k \quad (2)$$

$$\sum_k x_{kk} = P \quad (3)$$

$$\sum_k x_{ik} \le 1;\ \forall i \quad (4)$$

$$(c_{ik} + \alpha c_{kl} + c_{lj}) x_{ik} x_{jl} \le T;\ \forall i, j, k, l \quad (5)$$

$$x_{ik} \in \{0, 1\}\ \forall i, k \quad (6)$$

Expression (1), as the common objective of HMCPs, maximizes total utility of the hub network as sum of importance of the covered nodes. Constraint (2) ensures that non-hub nodes are only connected to hub nodes. Equation (3) guarantees that exactly P hubs are established in the network. Equation (4) confirms that a non-hub node may be connected to only one hub. Constraint (5) is suggested for covering an O/D pair. Regarding the first rule of covering, variables can simultaneously be equal to 1 only if the total transportation costs from origin i to hub k, hub k to hub l considering the discount factor and hub l to destination j are less than the given threshold T. To linearize the proposed non-linear model, binary variable y and constraint (7) are added to the model.

$$2y_{ij} \le \sum_k x_{ik} + \sum_l x_{jl}\ \forall i, j \quad (7)$$

$$y_{ij} \in \{0, 1\} \quad (8)$$

In constraint (7), binary variable y_{ij} is allowed to be 1 if both i and j are connected to the hub network simultaneously; otherwise, this variable is forced to be 0. Thereupon, non-linear objective function (1) can be substituted with (9).

$$\max \sum_i \sum_j w_{ij} y_{ij} \quad (9)$$

Using Lemma 1, constraint (5) can be linearized with no excess variables.

Lemma 1 Linear constraint (10) can be used instead of non-linear constraint (5).

$$(c_{ik} + \alpha c_{kl} + c_{lj}).(x_{ik} + x_{jl} - 1) \le T;\ \forall i, j, k, l \quad (10)$$

Proof Generally there are four possible cases for binary variables x_{ik}, x_{jl}. (1) $x_{ik} = x_{jl} = 0$: constraint (5) changes to $0 \le T$ and constraint (10) to $(-1)\left(c_{ik} + \alpha c_{kl} + c_{lj}\right) \le T$. T is a nonnegative parameter; so, both equations are obvious; (2) $x_{ik} = 0, x_{jl} = 1$: both equations change to $0 \le T$; (3) $x_{ik} = 1, x_{jl} = 0$: this case is similar to 2; (4) $x_{ik} = x_{jl} = 1$: both equations change to $\left(c_{ik} + \alpha c_{kl} + c_{lj}\right) \le T$. Since both constraints work similarly in all possible cases, we can use them interchangeably.

To the best of our knowledge, there are two distinct formulations for HMCPs presented in Campbell (1994) and Karimi and Bashiri (2011). The proposed non-linear model consists of n^2 variables and $n^4 + n^2 + n + 1$ constraints and the linearized version has $2n^2$ binary variables and $n^4 + 2n^2 + n + 1$ constraints. However, the formulation proposed by Campbell (1994) has $n^4 + n^2 + n$ variables with $2n^4 + n^2 + 1$ constraints and the model proposed by Karimi and Bashiri (2011) has $2n^2 + n$ variables with $n^4 + 2n^2 + n + 1$ constraints (using the same coverage type).

Multiple allocation HMCP model

To formulate multiple allocation HMCP, constraints (4) and (7) are omitted, and (11) and (12) are added.

$$y_{ij} \le \sum_{k} x_{ik} \; \forall i, j \tag{11}$$

$$y_{ij} \le \sum_{l} x_{jl} \; \forall i, j \tag{12}$$

These constraints allow upper bound of the binary variable to be 1 if both the origin and destination nodes are connected to hub network; otherwise, this equation is forced to be 0. This formulation for multiple allocation HMCPs has $2n^2$ binary variables and $n^4 + 3n^2 + 1$ constraints. Whereas the formulation proposed by Campbell (1994) has $n^4 + n^2 + n$ variables with $2n^4 + 1$ constraints and the model proposed by Karimi and Bashiri (2011) has $2n^2 + n$ variables with $n^4 + 2n^2 + n + 1$ constraints (using the same coverage type).

Bi-objective HMCPs

To increase safety of paths in the designed network, expressions (13), (14), (15) are added to proposed HMCP formulations.

$$\text{Max } s \tag{13}$$

$$s \le x_{ik} x_{jl} (p_{ik} p_{kl} p_{jl}) + M(1 - x_{ik} x_{jl}); \; \forall i, j, k, l \tag{14}$$

$$s \ge 0 \tag{15}$$

The second objective as expression (13) tries to maximize safety of the weakest path in the network considering the uncertainty in transmitting load to its destination. Constraint (14) provides an upper bound for safety of the weakest path in network. For O/D pair (i, j), the upper bound is equal to safety of established path i toward j via hubs k and l. Noteworthy, if the aforementioned path is not established, expression (14) is converted to an unnecessary constraint. Lemma 2 proposes a linear equivalent to non-linear constraint (14).

Lemma 2 Linear constraint (16) and non-linear constraint (14) may be used interchangeably.

$$s \le \frac{x_{ik} + x_{jl}}{2} p_{ik} p_{kl} p_{lj} + M\left(1 - \frac{x_{ik} + x_{jl}}{2}\right); \; \forall i, j, k, l \tag{16}$$

Proof There are four possible cases for binary variables x_{ik}, x_{jl}. (1) $x_{jl} = x_{ik} = 0$: right-hand side of both constraints equals M; (2) $x_{jl} = 1, x_{ik} = 0$: right-hand side of constraint (14) equals M and right side of constraint (16) equals $\frac{p_{ik} \cdot p_{kl} \cdot p_{jl}}{2} + \frac{M}{2}$ and considering M is a large number they work similarly; (3) $x_{jl} = 0, x_{ik} = 1$: this case is analyzed similar to 2; (4) $x_{jl} = x_{ik} = 1$: right-hand side of both equations equals $p_{ik} . p_{kl} . p_{jl}$. The two constraints work similarly in all possible cases so they might be used alternatively.

Solution algorithm: modified NSGA-II

In multi-objective models, existing constraints frequently prevent achieving a solution in which all the objective functions are optimal (Ghane and Tarokh 2012). In this situation, the set of Pareto optimal solutions, i.e. the solutions none of which is utterly better than the others, is the best choice. Classical optimization methods convert a given multi-objective problem to a single-objective one by different approaches. When the obtained single-objective problem is solved, in fact, one of the solutions in the set of Pareto optimal solutions is found. To map the whole Pareto optimal frontier, this procedure should be repeated many times, which is a time-consuming process (Deb et al. 2002). Furthermore, considering the complexity of real-world problems, a good approximation of Pareto optimal frontier is generally acceptable (Coello Coello 2007). This leads to the use of evolutionary algorithms to solve multi-objective problems. Early analogies between the mechanism of natural selection and learning or optimization process have been led to development of so-called evolutionary algorithms whose main goal is to simulate the evolutionary process on a computer (Coello Coello and Lamont 2004). A key benefit of multi-objective evolutionary algorithms is to attain a set of Pareto solutions. Such algorithms try to improve quality of the first-frontier members in consecutive generations. Initially, Schaffer (1985) applied a genetic algorithm (GA) to solve a multi-objective problem and proposed a vector-evaluated GA. After that, numerous multi-objective versions of evolutionary algorithms were proposed. Non-dominated sorting genetic algorithm (NSGA) is one of the most efficient and commonly used versions of multi-objective GA offered by Srinivas and Deb (1994). To resolve some shortcomings of NSGA, Deb et al. (2002) proposed an improved version called NSGA-II which in most multi-objective optimization problems is capable of converging to high-quality solutions with a better spread of solutions in the obtained frontier than the previous evolutionary algorithms (Noori-Darvish and Tavakkoli-Moghaddam 2012). In this paper, a modified version of NSGA-II is developed for solving the proposed multi-objective single allocation HMCP model. The modified NSGA-II differs in three ways from traditional one: (1) improved operators are designed to adapt to the problem, (2) an immigration operator is introduced for a better search in the solution space, and (3) a new mechanism is designed for adding individuals to the population. The following subsections are devoted to explain the proposed algorithm.

Chromosome structure

Besides simplicity, chromosome structure should contain all the information required to solve the problem. The most important features of problem are (a) ordinary nodes must be connected to hubs, (b) each ordinary node is allowed to connect only into one hub, (c) P hubs are established and (d) to cover an O/D pair, length of established path must be smaller than the specified covering radius.

In the proposed structure, each allele is denoted as a node, and number in it refers to the hub number to which the node is allocated (feature b). Hence, if a node's position is equal to the hub number allocated to it, the considered node is a hub. Figure 1 demonstrates the proposed structure for a problem with seven nodes. As alleles 2 and 5 are allocated to themselves, they are denoted as hubs. To create each chromosome in the initial population, P nodes are selected randomly to be hubs (feature c) and each one of the remaining nodes is assigned to only one hub randomly (feature a). Considering the covering constraint, for a specified O/D pair, if the total transportation cost in a given path is less than the determined covering radius it will be covered. For example, in the network designed in Fig. 1, ordinary nodes 4 and 6 are assigned to hubs 2 and 5, respectively. Thus, the considered path is 4–2–5–6. If the total transportation cost in this path $(c_{42} + \alpha c_{25} + c_{56})$ is less than the covering radius, O/D pair (4, 6) is covered and the associated utility (w_{ij}) is taken into account. Safety of the covered path 4–2–5–6 is $p_{42}p_{25}p_{56}$. This probability for the uncovered paths is not calculated.

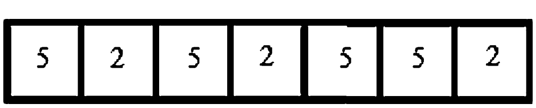


Fig. 1 Chromosome structure for a problem with seven nodes

Non-dominated sorting

Sorting and selecting the best individuals for the next generation are the most important differences between NSGA-II and the other multi-objective evolutionary algorithms. Initially, each solution is allocated a rank according to the number of times dominated by the other solutions. The rank of a solution determines the frontier in which it is located; therefore, the solutions with rank 1 are those in the first frontier. Actually, the first-frontier members are the algorithm's approximation of Pareto optimal frontier. Then, algorithm tries to improve them in iterative generations. The solutions located in a less crowded area are the more favorable solutions in the same frontier because of clarifying shape of the Pareto frontier.

To sort solutions of the same frontier, crowding distance measure is used (Deb et al. 2002). For each objective function, the solutions within the same frontier are sorted in ascending order. Based on the distance between two consecutive solutions for each objective function, the crowding distance measure is calculated as follows:

$$\mathrm{CD}_i = \sum_{m=1}^{M} \frac{f_m^{i+1} - f_m^{i-1}}{f_m^{\max} - f_m^{\min}} \quad (17)$$

where M is the number of objectives, $f_{\mathrm{m}}^{\max}$ and $f_{\mathrm{m}}^{\min}$ are the maximum and minimum amounts of objective m among frontier members, and $f_{\mathrm{m}}^{\mathrm{i}+1}$ and $f_{\mathrm{m}}^{\mathrm{i}-1}$ are the amounts of objective function m for the subsequent and precedent solutions in the sorted frontier population.

Due to the importance of boundary solutions of a frontier in detecting its shape, the first and the last solution crowding distances are set to be infinite. Finally, to select the best individuals for the next generation, the solutions are sorted in an ascending order according to their rank. Among the solutions with the same rank, those with higher crowding distance are preferred.

Genetic operators

Genetic operators are tools for better search, i.e. exploration and exploitation, in the solution space. As a result of mating, new offsprings are involved in population whose features are a mixture of parents' specifications. This process is simulated with a crossover operator in GA. Seldom abnormalities in the genetic structure of some individuals in the population cause salient differences in their specifications, called mutation operator in GA. The mutation operator plays an important role as it helps to escape local optima. Another phenomenon which human societies are faced is entrance of some foreign individuals into the existing population, called immigration, used as another genetic operator here.

Among various crossover techniques, we apply the simple single-point crossover. However, if we select a cut point randomly and interchange corresponding parts of chromosomes, infeasible solutions like those in Fig. 2 may be generated. As shown in Fig. 2, first offspring has four hubs including nodes 2, 3, 6, and 7 while second one has nodes 5 and 6 as hubs. Furthermore, in second offspring, the second and the forth nodes are allocated to non-hub nodes 7 and 2, respectively, which are as unacceptable allocations.

To resolve this problem, a modified single-point crossover is proposed. At first, the set of hub nodes in each parent, called h_1 and h_2, with cardinality of P are derived. Cut point Q is randomly selected from [1, $P-1$]. Then, the first Q members of h_1 and the last $P-Q$ members of h_2 are joined, and the set of hubs in first offspring c_1 is created. Similarly, the first Q members of h_2 and the last $P-Q$ members of h_1 make the set of hubs in second offspring c_2. If there are duplicates in each of the aforementioned sets, one of non-hub nodes is randomly replaced with one of repetitive hub nodes. After determining the set of hubs in each offspring, non-hub nodes are allocated to hubs randomly to create the complete offsprings. Figure 3 shows implementation of the proposed single-point crossover on Fig. 2.

To mutate a given chromosome, after random selection of a hub and non-hub node, non-hub node is altered to a hub and vice versa. Also, all the nodes allocated to the previous hub, including the hub node, are assigned to the new hub. To do mutation on the chromosome in Fig. 4, non-hub node 3 and hub node 2 are selected at random.

The third operator introduced here is called immigration. As it happens in most societies, a set of individuals, called immigrants, are added to the existing population periodically. Immigrants are created randomly and cause a better search in the solution space. In the proposed algorithm, set of immigrants consists of two parts. The first part is a fixed number of immigrants called 'basic number of immigrants' or BIN, and the second one varies in sequential iterations called 'variable number of immigrants' or VIN. After crossover and/or mutation, if the offspring dominates any of the parents, the operation is considered successful. The exact value of VIN is equal to number of unsuccessful crossover/mutation operations. In NSGA-II, results of two operators (crossover and mutation) are directly added to populations while, in the modified NSGA-II, only the successful offsprings will be added to population, and the others will be replaced with immigrants.

Accordingly, in each iteration of NSGA-II, number of offsprings and mutants added to population is more than (or equal to) the modified NSGA-II; however, sum of the offsprings, mutants, and immigrants in the modified NSGA-II is, as BIN value, more than sum of the offsprings and mutants in the traditional NSGA-II. The offsprings and immigrants are added to the main population, and after

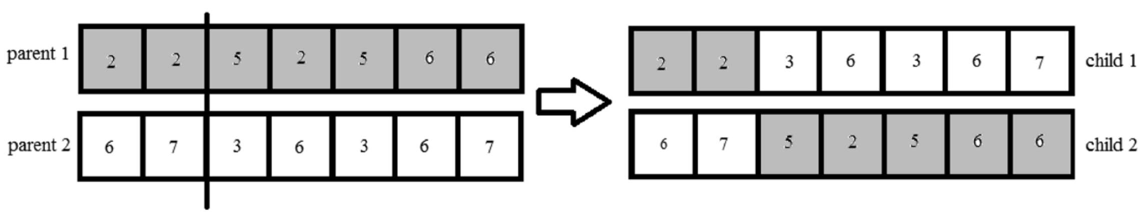
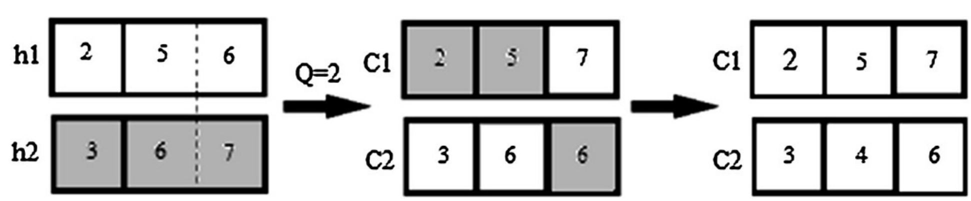


Fig. 2 Crossover with an infeasible solution

Fig. 3 Modified crossover performed on h1 and h2



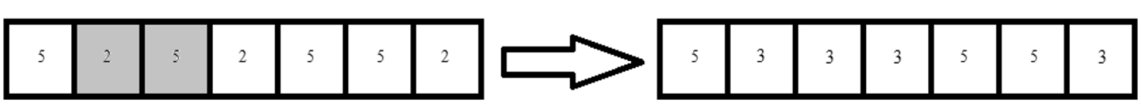


Fig. 4 A mutation example

using non-dominated sorting algorithm, better individuals, considering the population size, move to the next generation. The pseudo code of modified NSGA-II is provided in Fig. 5.

Computational results

In this section, first the proposed formulations are validated and their efficiencies are proved using some numerical examples extracted from the Turkish data set (Tan and Kara 2007) and effect of the second objective function is evaluated using a weighting method on a small-sized HMCP. Then, multi-objective metrics are introduced and after parameters setting, performance of the modified NSGA-II is studied versus the traditional algorithm.

Model validation and efficiency

In numerical examples, we have 10, 20, 30, 35 nodes and the number of established hubs are 1 and 2. Covering threshold (T) equals to average distance among nodes and discount factor α is 0.5. GAMS 22.2 with CPLEX solver is used to solve the experimental problems.

To validate the proposed linear formulations for single and multiple allocation HMCPs, they are compared to the formulations proposed in Karimi and Bashiri (2011) in terms of required computational time and relative gap (using the same coverage type). In Table 2, we compare the single allocation model with the one in Karimi and Bashiri (2011). In all experiments, the solution obtained by both formulations is the same which corroborates validity of the proposed model. The relative gap in the proposed

Fig. 5 Pseudo code of modified NSGA-II

```
1. Input parameters
   Population size (nPop), Crossover percentage (P_c), Mutation Percentage (P_m), Immigration
   percentage (P_img), Number of generations (Maxit), Offsprings' population size (Pop_c), Mutant's
   population size (Pop_m), Immigrants population size (Pop_img), Counter k for determining VIM
2. Initialization
   2.1. Create nPop individuals randomly and evaluate each one.
   2.2. Sort population using non-dominated sorting algorithm.
3. For iterations 1 to Maxit Do
   3.1. Let k=0
   3.2. Crossover
        3.2.1. Select P_c individuals using binary tournament selection and do Crossover.
        3.2.2. Evaluate offsprings.
        3.2.3. If any offspring dominates parents, add them to Pop_c, otherwise k=k+1.
   3.3. Mutation
        3.3.1. Select P_m individuals using binary tournament selection and do Mutation.
        3.3.2. Evaluate mutants.
        3.3.3. If mutant dominates parent, add it to Pop_m, otherwise k=k+1.
   3.4. Immigration
        3.4.1. Create P_img +k individuals randomly and add them to Pop_img.
        3.4.2. Evaluate Pop_img members.
   3.5. Add Pop_c, Pop_m, and Pop_img to the main population.
   3.6. Use non-dominated sorting algorithm to sort population.
   3.7. Move the first nPop individuals to the next generation.
   3.8. Store the first frontier members as Pareto frontier.
```

Table 2 Validating the proposed model for single allocation HMCP

Problem size	Number of Hubs	Proposed linear model		Proposed model in Karimi and Bashiri (2011)		Obtained solution
		Relative gap	Computational time (s)	Relative gap	Computational time (s)	
10	1	0.000	0.323	0.000	0.332	787,809
	2	0.000	0.362	0.000	0.405	1,270,931
20	1	0.063	2.767	0.091	2.872	1,777,083
	2	0.096	3.464	0.098	4.016	2,451,954
30	1	0.000	12.729	0.057	17.619	2,032,516
	2	0.092	16.632	0.099	24.616	2,746,645
35	1	0.089	26.636	0.099	33.887	6,333,382
	2	0.000	28.742	0.096	31.030	9,621,806

Table 3 Validating the proposed model for multiple allocation HMCPs

Problem size	Number of hubs	Proposed linear model		Proposed model in Karimi and Bashiri (2011)		Obtained solution
		Relative gap	Computational time (s)	Relative gap	Computational time (s)	
10	1	0.000	0.334	0.000	0.366	787,809
	2	0.000	0.367	0.060	0.376	1,286,698
20	1	0.000	2.772	0.000	2.812	1,777,083
	2	0.000	2.819	0.072	2.836	2,451,954
30	1	0.000	16.704	0.098	17.998	2,032,516
	2	0.088	22.820	0.097	25.092	2,918,431
35	1	0.000	30.067	0.095	33.312	6,333,382
	2	0.080	33.228	0.099	34.009	1,070,413

Table 4 Utility and safety probability for each O/D pair

Node	1	2	3	4	5	6	7
1	(0,1.000)	(261,0.262)	(28,0.424)	(178,0.730)	(187,0.521)	(72,0.988)	(59,0.335)
2	(261,0.262)	(0,1.000)	(326,0.508)	(179,0.489)	(123,0.232)	(105,0.038)	(79,0.680)
3	(28,0.424)	(326,0.508)	(0,1.000)	(286,0.579)	(329,0.489)	(165,0.885)	(152,0.137)
4	(178,0.730)	(179,0.489)	(286,0.579)	(0,1.000)	(307,0.624)	(80,0.0.913)	(109,0.721)
5	(187,0.521)	(123,0.232)	(329,0.489)	(307,0.624)	(0,1.000)	(296,0.796)	(324,0.107)
6	(72,0.988)	(105,0.038)	(165,0.885)	(80,0.0.913)	(296,0.796)	(0,1.000)	(151,0.654)
7	(59,0.335)	(79,0.680)	(152,0.137)	(109,0.721)	(324,0.107)	(151,0.654)	(0,1.000)

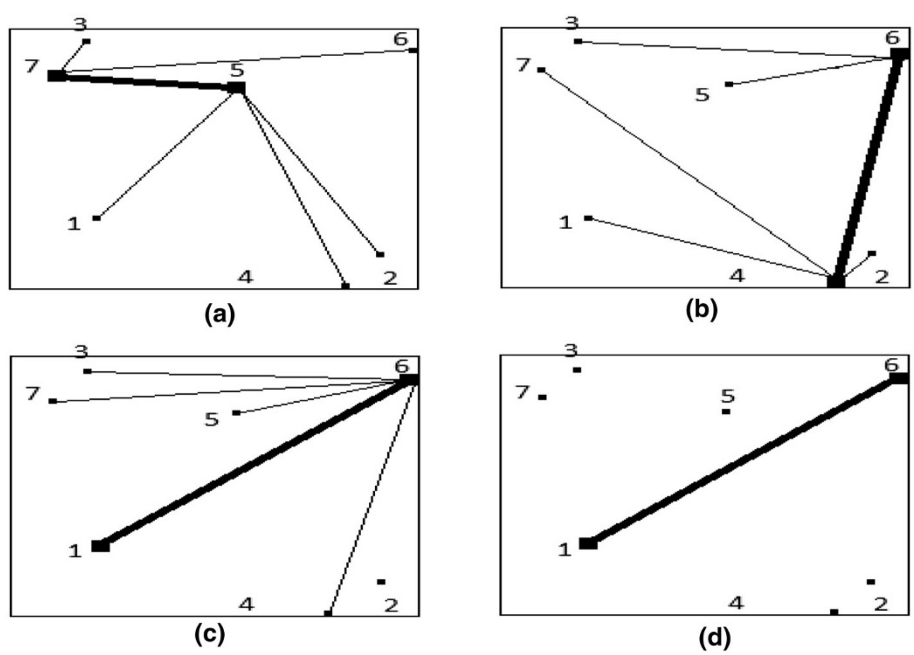


Fig. 6 Optimal results of multi-objective model for different weighting preferences

model is always less than that in Karimi and Bashiri (2011). Average relative gap for the proposed model is 0.042 whereas in Karimi and Bashiri (2011) it is 0.067. Also, computational time for the proposed model is less than that of Karimi and Bashiri (2011) for all experimental problems.

Similar experiments for the proposed model in multiple allocation case result to the same solutions for both formulations. Table 3 shows that average relative gap for the proposed model is 0.021 whereas in Karimi and Bashiri (2011) it is 0.065. Also, computational time in all experiments is always less than that in Karimi and Bashiri (2011).

Effects of the second objective

We involve a second objective to force the model to choose safer paths for connecting all O/D pairs. To represent optimum results of the proposed multi-objective model for a small-sized instance, a weighting method is applied for obtaining Pareto solutions. Consider a problem with seven nodes in which two hubs must be established. The covering radius and discount factor are assumed to be 300 and 0.5, respectively. Table 4 presents utility of covering and safety probability of each O/D pair. Each time a given weight is allocated to each objective, and GAMS22.2 is used to solve the resulting single-objective problem.

Figure 6 demonstrates the optimal solutions from four different weighting preferences. Case (a) shows optimal solution when the only criterion is maximizing the total utility ($w_1 = 1, w_2 = 0$). In this case, nodes 5 and 7 are hubs, total covering utility is 3,796 and the path from 2 to 3 via hubs 5 and 6 is the weakest one with safety 0.003. In case (b), the associated weights for both objectives are 0.5 ($w_1 = w_2 = 0.5$). Accordingly, nodes 4 and 6 are hubs, total covering utility is 3,796 and the path from 2 to 7 via hub 4 is the weakest one with safety 0.353. In case (c), $w_1 = 0.4$ and $w_2 = 0.6$. As a result, total covering utility is 2,723, node 2 is not connected to network, and the weakest path is the one from 5 to 7 via hub 6 with safety 0.521. Finally, in case (d), the only criterion is maximizing safety of weakest path ($w_1 = 0, w_2 = 1$). Consequently, nodes 5 and 6 are hubs with safety 0.988, the other nodes are not connected to network, and total covering utility is 72.

The above results clearly show effects of the second objective on forming the hub network (i.e. selecting hub nodes and paths for linking non-hub nodes). Obviously, an increase in the importance of second objective causes selection of more reliable paths although the total covering utility may be decreased.

Multi-objective metrics

Quality of solutions and their relative dispersion in Pareto frontier are the most important properties of an evolutionary algorithm. To compare modified versus traditional NSGA-II, five multi-objective metrics are introduced.

Table 5 Desired levels for parameters and selected values

Parameters	Factor levels				Selected value
	1	2	3	4	
Maximum number of iterations (MaxIt)	60	70	100	150	70
Population size (*n*Pop)	80	100	150	200	100
Crossover rate (P_c)	0.7	0.75	0.8	0.85	0.7
Mutation rate (P_m)	0.05	0.10	0.15	0.20	0.2
BIN	0.2	0.3	0.4	0.5	0.4

Table 6 L_{16} design and related response levels

Experiment	Parameters					Average response level
	MaxIt	*n*Pop	P_c	P_m	BIN	
1	60	80	0.7	0.05	0.2	17.5214
2	60	100	0.75	0.10	0.3	31.9664
3	60	150	0.8	0.15	0.4	7.76250
4	60	200	0.85	0.20	0.5	11.8583
5	70	80	0.75	0.15	0.5	16.4317
6	70	100	0.7	0.20	0.4	39.7856
7	70	150	0.85	0.05	0.3	24.7537
8	70	200	0.8	0.10	0.2	22.3217
9	100	80	0.8	0.20	0.3	19.4879
10	100	100	0.85	0.15	0.2	11.4221
11	100	150	0.7	0.10	0.5	9.24550
12	100	200	0.75	0.05	0.4	19.8328
13	150	80	0.85	0.10	0.4	17.5379
14	150	100	0.8	0.05	0.5	17.7215
15	150	150	0.75	0.20	0.2	21.6488
16	150	200	0.7	0.15	0.3	24.8538

1. *Quality metric (QM)* More solutions in Pareto frontier imply better performance of the algorithm (Schaffer 1985).
2. *Best frontier members (BFM)* Solutions with the best fitness for each objective in Pareto frontier.
3. *Average frontier fitness (AFF)* Average fitness of solutions in first frontier for each objective.
4. *Mean ideal distance (MID)* Average distance among solutions in Pareto frontier and a hypothetical ideal solution (Zitzler and Thiele 1998). Lower value of MID shows better performance of the algorithm.

$$\mathrm{MID} = \frac{\sum_{i=1}^{n} \sqrt{\left(\frac{f_{1i} - f_1^{\mathrm{best}}}{f_{1,g}^{\max} - f_{1,g}^{\min}}\right)^2 + \left(\frac{f_{2i} - f_2^{\mathrm{best}}}{f_{2,g}^{\max} - f_{2,g}^{\min}}\right)^2}}{n} \quad (18)$$

where n is the number of Pareto solutions, f_{ji} is value of jth objective for ith solution in Pareto frontier, and

Table 7 Computational results for the QM_1 and QM_2

Size	BFM				AFF			
	NSGA-II		Modified NSGA-II		NSGA-II		Modified NSGA-II	
	First objective	Second objective	First objective	Second objective	First objective	Second objective	First objective	Second objective
20	7,617	0.9708	7,654	0.9732	4,489.6	0.9151	4,640.3	0.9152
40	14,201	0.9236	17,353	0.9367	9,577.0	0.8933	10,048.4	0.8955
50	33,244	0.9164	35,117	0.9251	16,746.2	0.8593	16,981.9	0.8593
70	47,767	0.9006	49,273	0.9047	28,968.2	0.8459	29,709.5	0.8483
100	97,469	0.8702	104,518	0.8703	68,153.9	0.8357	68,577.3	0.8360
150	24,407	0.9292	30,317	0.9509	14,194.7	0.8956	15,014.2	0.8992
200	72,443	0.9035	86,538	0.9035	59,677.1	0.8728	61,254.3	0.8834
300	65,970	0.9042	72,630	0.9050	52,037.9	0.8714	55,272.2	0.8735
400	145,856	0.9012	151,884	0.9014	12,807.7	0.8594	126,212.7	0.8638
500	124,318	0.9016	134,912	0.9017	109,814.7	0.8766	114,249.5	0.8875
1,000	372,591	0.9002	375,304	0.9002	345,237.9	0.8675	347,650.6	0.8679

$f_{i,g}^{\max}$, $f_{i,g}^{\min}$ are maximum and minimum amounts of ith objective among solutions in Pareto frontier. Notably, coordinates of the ideal solution in all problems are assumed to be $(f_1^{\text{best}}, f_2^{\text{best}}) = (\sum_{i,j} w_{ij}, 1)$.

5. *Spacing metric (SM)* Distribution of solutions in Pareto frontier as denoted in (19). Lower SM signs better dispersion of the solutions in the frontier (Srinivas and Deb 1994).

$$\text{SM} = \frac{\sum_{i=1}^{n-1} \left| d_i - \bar{d} \right|}{(n-1)\bar{d}} \quad (19)$$

where d_i is Euclidean distance between solutions i and $i+1$ in sorted Pareto solutions and $\bar{d}$ is average Euclidean distance.

Parameter tuning

Parameter tuning has a salient effect on quality of solutions and computational time of evolutionary algorithms. To set parameters, Taguchi method is applied which is a fractional factorial experiment design. Compared to the full factorial experiments, fractional approaches focus on the orthogonal arrays of data to analyze different levels of factors which lead to a considerable reduction in number of conducted experiments. For a comprehensive review on the Taguchi approach, one may refer to Roy (2001).

As mentioned earlier, speed of convergence to actual Pareto frontier and diversity of solutions in obtained frontier are the main criteria in analyzing evolutionary algorithms. Rahmati et al (2013) incorporate these parameters and propose a single response metric for Taguchi method namely multi-objective coefficient of variation (MOCV).

$$\text{MOCV} = \frac{\text{QM}}{\text{SM}} \quad (20)$$

There are five parameters in NSGA-II that need to be tuned, each of which is assigned four initial values as Table 5 based on the previous experiments. At last, the mentioned factors are analyzed with an L_{16} design. Larger values of response are more desirable. Each experiment is done four times and average response levels are given in Table 6. Given the response levels, parameter levels are provided in Table 5.

Performance of modified NSGA-II

To assess performance of the proposed NSGA-II, it is compared with the traditional one. They are compared regarding the multi-objective metrics introduced in 4.3. Generally, hub location problems are among NP-hard ones and they have a high computational complexity. Due to the size of numerical examples in other papers (for a comprehensive list of instances, one may refer to Tables 1, 2 in (Farahani et al. 2013)), we have divided our experiments into small, medium, and large-sized problems. Small-size instances have 20, 40, and 50 nodes, medium-sized problems have 70, 100, 150, and 200 nodes, and the large-sized ones consist of 300, 400, 500 and 1,000 nodes. The experiments were done on a PC with Core 2 Duo, CPU2.4 GHs and RAM 1 GB. MATLAB R2011b was used to code both the traditional and modified NSGA-II. The number of hubs in each experimental problem is randomly selected from the number of nodes. The nodes are scattered on a plane according to problem size, and Euclidian distance is calculated for them. Hence, the distances among the nodes satisfy triangle inequality. For problems of less than 100 nodes, between 100 and 500 nodes, and more than 500

Table 8 Computational results for QM, SM and MID metrics

Size	QM		SM		MID	
	NSGA-II	Modified NSGA-II	NSGA-I	Modified NSGA-II	NSGA-II	Modified NSGA-II
20	17	25	0.8087	0.7147	4.43	2.64
40	10	15	0.8728	0.5654	10.56	9.75
50	22	25	0.8121	0.7631	7.31	6.76
70	19	24	0.8166	0.6917	10.22	10.53
100	14	18	0.6987	0.5454	15.37	11.61
150	17	22	0.6693	0.6691	92.39	69.64
200	13	19	0.9695	0.8631	82.77	59.61
300	16	18	0.8517	0.7740	16.81	16.93
400	11	16	1.0470	0.8465	13.29	9.65
500	12	12	0.7820	0.7612	52.33	35.11
1,000	14	16	0.9158	0.8660	62.88	46.71

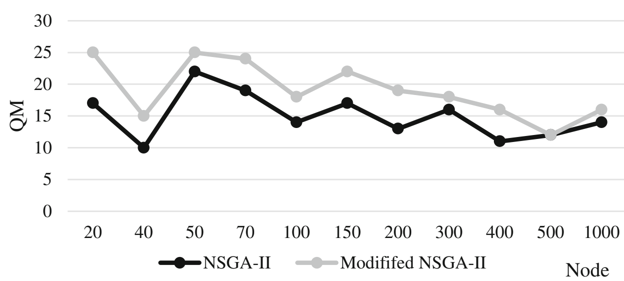


Fig. 7 QM in NSGA-II versus modified NSGA-II

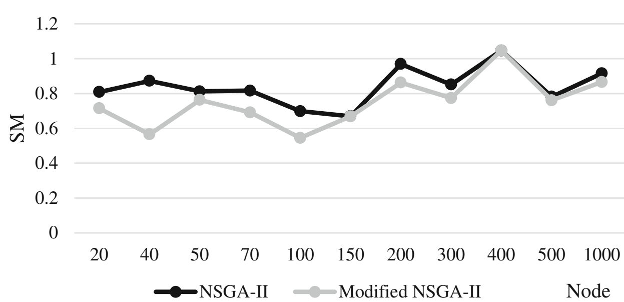


Fig. 8 SM in NSGA-II versus modified NSGA-II

nodes, coordinates of the network planes are selected at random from [0,100], [0,300], and [0,500], respectively. Also, the distance matrix is symmetric. Covering each node in HMCPs creates a certain utility according to some properties of that node. Utility is selected randomly from [0,350]. Comparative results on traditional and modified NSGA-II algorithms for all multi-objective metrics are provided in Tables 7 and 8.

It is obvious that modified NSGA-II surpasses the traditional one according to BFM metric in all experimental problems. For first objective, the most deviation of traditional NSGA-II from modified one is 24.21 % for problems of size 150, and the average deviation is 9.58 %. For

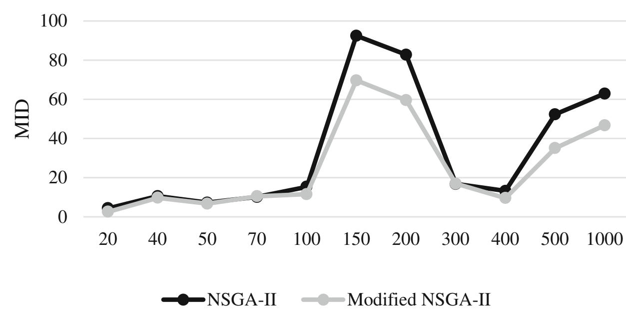


Fig. 9 MID in NSGA-II versus modified NSGA-II

Table 9 ANOVA results in 5 % risk level

Metric	F statistic	P_{value}	Decision on null hypothesis
QM	5.7365	0.0271	Rejected
SM	5.2741	0.0348	Rejected
MID	0.4474	0.5052	Not enough evidence to reject

second objective, corresponding most deviation is 2.29 % for problems of size 150, and the average is 0.5 %.

AFF in modified NSGA-II surpasses that of traditional one which shows better performance of proposed algorithm. Considering Table 7 for the first objective, the most and mean AFF deviation of traditional NSGA-II from the modified one are 5.85 % (problem of size 300) and 3.2 %, respectively. The corresponding deviations for second objective are 1.23 % (problem of size 500) and 0.38 %, respectively.

Table 8 summarizes the values of QM, SM and MID in experimental problems. As mentioned earlier, larger values for QM and smaller values for SM and MID are more desirable. In all experimental problems, QM and SM in modified NSGA-II are better than those in traditional algorithm. This is also evident in Figs. 7 and 8. For MID

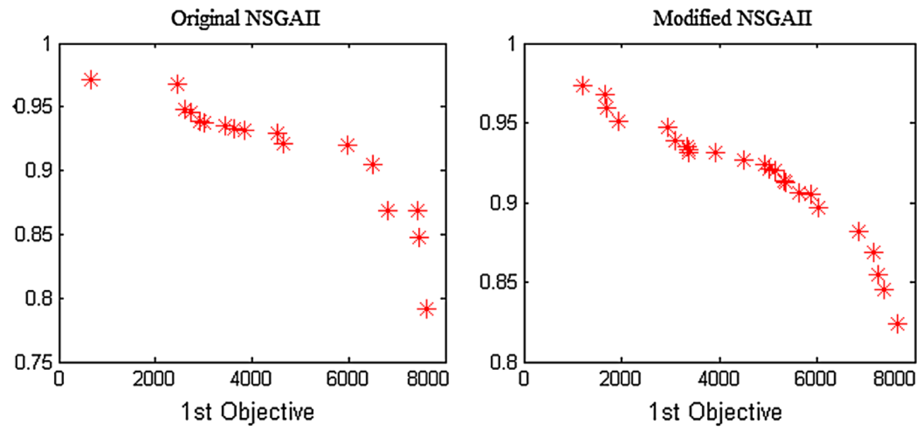


Fig. 10 Pareto frontier of problem with 20 nodes

metric, except problems with 100 and 300 nodes, modified NSGA-II shows a better performance than NSGA-II. Figure 9 compares MID metric of NSGA-II with that of modified NSGA-II.

Results of ANOVA established for two algorithms in terms of QM, SM and MID are provided in Table 9. Null hypothesis is that there is no significant deviation among the three metric values of two algorithms. According to Table 9 with a 5 % significance level, null hypothesis is rejected for QM and SM which means that modified NSGA-II outperforms NSGA-II. Although in most cases MID in the proposed algorithm is less than the original one, but the respective gap is not significant. As a sample, Pareto frontiers of problem with 20 nodes for both algorithms are displayed in Fig. 10. Horizontal and vertical axes stand for first and second objectives, respectively. The diagram confirms the monotony and higher quality of Pareto frontier of the modified NSGA-II.

Concluding remarks and directions for future research

In this paper, we proposed new formulations for single and multiple allocation hub maximal covering problems. It was shown that proposed non-linear model and linearized versions outperform the existing formulations from the literature. Also, the considered problems were investigated under uncertainty in transmitted loads to destinations via a bi-objective mixed-integer model. Such models are applied in martial transportation networks, message delivery in telecommunication systems and air transport systems. Along with maximization of coverage, selecting safer paths for transmitting loads was considered as another objective. To solve the proposed model, a modified version of NSGA-II was developed in which new crossover and mutation operators are introduced to adapt with structure of the considered problem. Moreover, a new immigration operator was involved. The modified NSGA-II and NSGA-II were compared using five multi-objective metrics, four of which proved the supremacy of the proposed algorithm. In this paper, the probability of disruption in a link was investigated; however, an interesting direction for future research is addressing probability of disruption for nodes.

References

Alumur SA, Kara BY, Karasan OE (2012a) Multimodal hub location and hub network design. Omega 40(6):927–939

Alumur SA, Nickel S, Saldanha-da-Gama F (2012b) Hub location under uncertainty. Transp Res Part B Methodol 46(4):529–543

Calık H, Alumur SA, Kara BY, Karasanet OE (2009) A tabu-search based heuristic for the hub covering problem over incomplete hub networks. Comput Oper Res 36(12):3088–3096

Campbell JF (1994) Integer programming formulations of discrete hub location problems. Eur J Oper Res 72(2):387–405

Campbell JF, O'Kelly ME (2012) Twenty-five years of hub location research. Transp Sci 46(2):153–169

Coello Coello CA, Lamont GB (2004) Applications of Multi-Objective Evolutionary Algorithms Advances in Natural Computation, In: Yao X (ed) vol 1, World Scientific Publishing Co, Singapore

Coello Coello CA, Lamont G, van Veldhuizen DA (2007) Evolutionary algorithms for solving multi-objective problems (genetic and evolutionary computation). 2nd edn, In: Goldberg, DE, Koza JR (eds) Springer, Heidelberg, p 810

Contreras I, Cordeau J-F, Laporte G (2011) Stochastic uncapacitated hub location. Eur J Oper Res 212(3):518–528

Davari S, Fazel Zarandi MH, Turksen IB (2013) The incomplete hub-covering location problem considering imprecise location of demands. Sci Iran 20(3):983–991

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. Evol Comput IEEE Trans 6(2):182–197

Eydi A, Mirakhorli A (2012) An extended model for the uncapacitated single allocation hub covering problem in a fuzzy environment. In: Proceedings of the international multiconference of engineers and computer scientists, International Association of Engineers, Hong Kong

Farahani RZ, Hekmatfar M, Boloori A, Nikbakhsh E (2013) Hub location problems: a review of models, classification, solution techniques, and applications. Comput Ind Eng 64(4):1096–1109

Fazel Zarandi MH, Davari S, Haddad Sisakht SA (2012) The Q-coverage multiple allocation hub covering problem with mandatory dispersion. Sci Iran 19(3):902–911

Ghane M, Tarokh MJ (2012) Multi-objective design of fuzzy logic controller in supply chain. J Ind Eng Int 8(1):1–8

Hwang YH, Lee YH (2012) Uncapacitated single allocation p-hub maximal covering problem. Comput Ind Eng 63(2):382–389

Kara BY, Tansel BC (2003) The single-assignment hub covering problem: models and linearizations. J Oper Res Soc 54(1):59–64

Karimi H, Bashiri M (2011) Hub covering location problems with different coverage types. Sci Iran 18(6):1571–1578

Mohammadi M, Jolai F, Rostami H (2011) An M/M/c queue model for hub covering location problem. Math Comput Model 54(11–12):2623–2638

Mohammadi M, Jolai F, Tavakkoli-Moghaddam R (2013) Solving a new stochastic multi-mode p-hub covering location problem considering risk by a novel multi-objective algorithm. Appl Math Model 37(24):10053–10073

Noori-Darvish S, Tavakkoli-Moghaddam R (2012) Minimizing the total tardiness and make span in an open shop scheduling problem with sequence-dependent setup times. J Ind Eng Int 8(1):1–13

O'Kelly ME (1986) The location of interacting hub. Transp Sci 20(2):92–106

O'Kelly ME (1987) A quadratic integer program for the location of interacting hub facilities. Eur J Oper Res 32(3):393–404

Qu B, Weng K (2009) Path relinking approach for multiple allocation hub maximal covering problem. Comput Math Appl 57(11–12):1890–1894

Rahmati SHA, Hajipour V, Akhavan Niaki ST (2013) A soft-computing Pareto-based meta-heuristic algorithm for a multi-objective multi-server facility location problem. Appl Soft Comput 13:1728–1740

Roy RK (2001) Design of experiments using the Taguchi approach: 16 Steps to product and process improvement, vol 1. Wiley, New York, p 560

Schaffer JD (1985) Multiple objective optimization with vector evaluated genetic algorithms, in proceedings of the 1st international conference on genetic algorithms, L. Erlbaum Associates Inc, Hillsdale, pp 93–100

Sim T, Lowe TJ, Thomas BW (2009) The stochastic-hub center problem with service-level constraints. Comput Oper Res 36(12):3166–3177

Srinivas N, Deb K (1994) Multiobjective optimization using nondominated sorting in genetic algorithms. Evol Comput 2(3):221–248

Tan PZ, Kara BY (2007) A hub covering model for cargo delivery systems. Networks 49(1):28–39

Tavakkoli-Moghaddam R, Forouzanfar F, Ebrahimnejad S (2013) Incorporating location, routing, and inventory decisions in a bi-objective supply chain design problem with risk-pooling. J Ind Eng Int 9(1):1–6

Wagner B (2008) Model formulations for hub covering problems. J Oper Res Soc 59(7):932–938

Yang T-H (2009) Stochastic air freight hub location and flight routes planning. Appl Math Model 33(12):4424–4430

Zhai H, Liu Y, Chen W (2012) Applying minimum-risk criterion to stochastic hub location problems. Proc Eng 29:2313–2321

Zitzler E, Thiele L (1998) Multiobjective optimization using evolutionary algorithms—a comparative case study, in parallel problem solving from nature—PPSN V. In: Eiben A et al (eds) Springer, Heidelberg, pp 292–301