Research in Transportation Economics 90 (2021) 100896

Contents lists available at ScienceDirect

Research in Transportation Economics

journal homepage: http://www.elsevier.com/locate/retrec

A Bi objective uncapacitated multiple allocation p-hub median problem in public administration considering economies of scales

Aliasghar Tofighian [a], Alireza Arshadi khamseh [b,*]

[a] Department of Industrial Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran
[b] Department of International Logistics and Transportation, Faculty of Economics Administration and Social Sciences, Istanbul Gelisim University, Istanbul, Turkey

ARTICLE INFO

Classification codes:
RCGHP

Keywords:
P-Hub median
Imperialist competitive algorithm
Competitive location
Public administration transportation
Economics of scale

ABSTRACT

This paper addresses uncapacitated multiple allocation p-hub median problems, which deals with both the constructors' and the users' objectives in order to obtain an economically sustainable system. One objective is maximizing the overall investment return in road and hub construction and the users' satisfaction is translated by minimization of the overall usage cost. The problem is formulated in a way that can cover three possible policies as: Governmental requirement, constructor's break-even point and predefined make span. To make these models more pragmatic, variable discount factors are used in preference to fixed ones. Accordingly, a comprehensive discussion about discount factors and their components has been included to justify the use of variable discount factors. Then some meta-heuristic algorithms like the Imperialist competitive algorithm (ICA), and an enhanced variation of a well-known multi-objective genetic algorithm called nondominated sorting genetic algorithm II (NSGA-II) are developed and applied to solve the problem. The performance of algorithms is compared to each other by utilizing some indicators such as hypervolume, ε-indicator, spacing metric, and CPU time. Computational experiments emphasize the need for using stated assumptions and the variable discount factor. It also confirms the efficiency of the proposed ICA.

1. Introduction

In the last two decades, the use of the hub transportation systems has aroused. These systems are able to facilitate the flow of entities by offering the possibility of efficient capacity sharing and fleet management and can be utilized to transfer entities in the telecommunication, transportation, and logistics systems (Gelareh & Nickel, 2011). Hub location problems (HLPs) are defined as a network in which there are n interacting points called spokes (non-hub nodes) and p centers of transportation named hubs (Damgacioglu H., Dinler, Ozdemirel, & Iyigun, 2014) and each pair of origin/destination are linked through at least one hub facility. One of the main characteristics of hub systems is that, by gathering the flows, they enable economies of scale to result in lower transportation costs (Klincewicz, 2002). The p-hub median problem can be defined as follows: given a set of nodes with pairwise traffic demands, choose p nodes to locate hubs and route the traffic of all nodes through these hubs at minimum cost. It is often assumed that hubs are connected by a complete network, the routing cost between hubs is discounted at a factor $0 < \alpha \leq 1$, and no direct connection exists between two none hub nodes (Yaman, 2011).

Despite some previous researches e.g. (O'Kelly, Morton, & Harvey, 1994) and (Kimms, 2005), the impossibility of direct linking between each pair of origin/destination (Martí, Corberán, & Peiró, 2015) is one of the classic assumptions and simplifications of HLPs which almost true in telecommunication systems but not in transportation systems, because in real-world cases, there are already several direct links between some pairs of nods and customers are able to use them instead of an indirect system, especially in air transportation, direct shipment is usually one of the options. Therefore, in this study the available direct routes and also any other existing systems are accessible for users and may reduce the share and the profits of the proposed hub system. Most studies in the literature attempted to model a hub network as the only available transportation system. However, in real-world cases, there are already some other transportation systems and competitive environments usually overcast establishing a new hub location system (Luer-Villagra & Marianov, 2013; Sasaki, Campbell, Krishnamoorthy, & Ernst, 2014). For all we know, nearly all previous studies considered minimizing all costs as the one objective without considering any differences between costs of the system and its earned benefits, even though the origin of costs and gained benefits are not the same. Note that in reality,

* Corresponding author.
E-mail addresses: aatofighian@gmail.com (A. Tofighian), akhamseh@gelisim.edu.tr (A. Arshadi khamseh).

https://doi.org/10.1016/j.retrec.2020.100896
Received 18 May 2019; Received in revised form 27 May 2020; Accepted 3 June 2020
Available online 12 August 2020

the system that reduces users' costs more than the others isn't the one that is grounded. Because almost all transportation systems are established by constructors (mainly governments), not by its users, and these institutions don't establish the system simply to reduce users' costs.

In constructors' sight, an appealing system is a profitable system that at least can compensate for their expenses. Therefore, studying and treating all types of costs, in the same way, isn't an apt way to approach and model these cases. For instance, suppose that by establishing a hub system, public cost (customers' cost) is reduced by X and this system costs Y for constructors, so if we study X and Y, in the same way, the objective function should be minimizing Y-X. In reality, however, up until constructors are not justified about how Y should be compensated, they don't invest in the system.

Furthermore, in reality, there is not just one option for users. They don't necessarily have to use the system that constructors have established. All users have the choice to utilize the cheapest and most cost-effective shipping plan after checking almost all available options. As a tangible example in postal systems, and for close distances users tend to send their cargo directly using intercity delivery options (Uber for example), instead of a public posting system, because it is cheaper, faster and needs less paperwork. In this atmosphere, it is clear for constructors that they will be unable to gain all the shipping shares. However, in most of the previous research, it is assumed that all shipping between all origin-destination pairs should be fulfilled by the proposed system. According to these facts and their preferred time window for the return of investments, constructors make a system that maximizes their revenue (and not necessarily a system that minimizes users' costs); they are also aware that the more users use the system the more revenue they get. Therefore, the problem is to find the best and attractive system for constructors and beneficial for users simultaneously.

Another flaw in previous studies is the lack of comprehensive study about discount factors that play a very important role in constructors and users' decision making. However, they are considered to be fixed (for example, 10–50%) or at the best, they are considered to be dependent on distance or amount of the shipload. In other terms the more distance and/or cargo the more discount, but it is not always the case. For example, for valuable assets like gold, fragile products, military equipment, nuclear waste, etc. these discount factors are negatively affected by distance and amount of cargo. Note that some materials like nuclear wastes and military equipment, don't even have the potential to be stored and shipped massively together. Some other types of goods require insurance due to their characteristics and associated charges also increase the cost of the proposed system. To cover the mentioned gaps in the literature and tackle these issues, a comprehensive discussion about discount factors is presented, in later sections.

In summary, in this research instead of using total cost as the only objective function, we separate costs and benefits by their payers and beneficiaries. Because, constructors want to maximize their own profit not the users' costs and if a system is not interesting enough for them they simply don't establish it at all. In contrast the users have no obligations to use the established system and they use a system that minimizes their costs. The problem is formulated in a way that can cover three possible policies with some minor changes: 1) constructors (generally governments) want to establish a system to facilitate the public transportation even if the system is not economically feasible for them. In that case, earning financial benefits is not a high priority, 2) constructors intended to participate in an economically sustainable project and achieve the break-even point within the very first period, and 3) constructors want to participate in an economically feasible project and gain break-even point within a predefined makespan. In any of them, constructors are fully aware that in a competitive environment they can't take over and cover all the customers' shipping, therefore they try to maximize their own profit and make the system more attractive for users by reducing their costs. Therefore, a mathematical model that considers both the viewpoints of users and constructors is formulated that focuses on establishing a hub network wherein customers can choose whether to use it or not and also covers constructors' benefits. Then some meta-heuristic methods are proposed to solve the models and examine their effectiveness. Furthermore, the different aspects of discount factors and their root causes such as fuel consumption, maintaining, and repair costs, assurance and holding costs, taxes and tolls, etc. are discussed.

The rest of the paper is organized as follows: in the next section a bi-objective mathematical model concerning mentioned policies is presented (an illustrative example is presented to describe the model in practice see APPENDIX B). Then different aspects of discount factors are discussed section 3. A new hybrid solution method based on the imperialist competitive algorithm is introduced to solve the model in section 4. Results and computational experiences are summarized in section 5. Section 6 comprises conclusions and directions for further research.

2. Mathematical model

In this section, a model based on (Campbell, 1996) is formulated. Let N be the number of all nodes that exchanges traffic and potentially at each of these nodes, a hub facility can be established. For any pair of origin/destination nodes (Non hub nodes) i and j ($i,j \in N$) there is an $h_{i,j}$ flow of cargo (amount of the cargo), which must be shipped and the distance between each pair of nodes is given as $d_{i,j}$. Consider that the established hub system is not the only option and customers can ship their cargo through the hub system or other available systems (in this study direct shipment is one of the options). Shipping cost between each pair of nodes i,j is $C_{i,j}$, the cost of shipping through the best available system is predefined and denoted as $CC_{i,j}$ and the shipping cost through hub nodes k and m ($k,m \in N$) is $C_{i,j}^{k,m}$. It's worth mentioning that when a direct shipment is the only option $CC_{i,j} = C_{i,j}$. In this competitive atmosphere, a constructor wants to invest a limited budget (HB) to establish P hub facilities (choose P nodes as hub nodes). On one hand, this may bring some fixed costs like establishing costs on hub nodes (hb_k) and also some annual maintenance costs ($HC_{k,m}$), but on the other hand, the constructor can earn benefits by charging users with tax/toll ($T_{k,m}$) for using the systems. The tax rate ($T_{k,m}$) can be applied on the total amount of cargo that flows between each two hub nodes k and m. However, if constructors provide services for collecting and distributing cargo, they can charge users for their service from the origin to the first hub ($T_{i,k}$) as well as from the second hub to the destination ($T_{m,j}$). This study considers inter hub taxes ($T_{k,m}$) as the only source of income for the constructors. Note that maintenance costs $MC_{k,m}$ and $T_{k,m}$ are annual costs and benefits of the constructors and in order to homogenize them with initial costs constructors should calculate their net present values (NPV) using the rate of return (RR). It is worth mentioning that the rate of return is the net gain or loss on an investment over a specified time period, expressed as a percentage and usually the minimum RR is a bank interest rate (2%–18% based on the country and the banking policies). The users can benefit from three types of discount factors e.g. collection cost factor (discounts between origin and the first hub node), transfer cost factor (discounts between hub nodes), and distribution cost factor (discounts between the second hub and destination) which are indicated as $\chi_{i,k}$, $\alpha_{k,m}$, and $\sigma_{m,j}$ respectively. In order to model this problem two binary decision variables are needed, first X_k that gets one if a hub established in node k, otherwise gets zero, and the second is $Z_{i,j}^{k,m}$ that gets one if cargo delivered from origin i to destination j through hub k and m, otherwise gets zero. Moreover, the net benefit of constructors ($\mathfrak{B}$) and the total cost of users ($\mathfrak{C}$) can be calculated using mentioned parameters and these two binary decision variables. Note that if constructors want to gain the break-even point within a makespan of MS, the model needs an integer decision variable for the time as t, and it is obvious that t should be more than one year and less than the makespan (MS).

In the following, characteristics and assumptions of the problem determine the scope of the current study:

1. Each node can link to all hub nodes (multiple allocations)
2. The number of hubs is determined and equals to P.
3. Each hub facility can afford an unlimited flow of shipments (unlimited capacity).
4. The cost of implementing a new facility is determined and fixed (fixed cost).
5. The available budget is limited.
6. Competing systems are available. At least a direct shipment system is available on some routes.
7. Taxes are fixed and much lower than discounts.
8. Paths between hubs need annual maintenance.
9. Maintenance costs are fixed for each path.
10. The inter hub (hub-to-hub) graph is complete.

Note that some parameters and variables like hb_k are presented with more than one character, so hb_k do not indicate $h \times b_k$. We used "." or " $\times$ " as the multiplication mark in the mathematical formulations. In the following, we discuss three possible policies that constructors may choose.

2.1. 1. Policy 1: facilitating transportation system

Within this policy, public benefit is the top priority so constructor (government in this case) merely wants to establish a hub system to facilitate transportations and only if it is possible constructor tries to gain some financial benefits as well.

$$\max \mathfrak{B} \text{ and } min \ \mathfrak{C} \tag{1}$$

subject to:

$$\sum_k x_k = P \tag{2}$$

$$\sum_k x_k.hb_k \le HB \tag{3}$$

$$\sum_k \sum_m Z_{i,j}^{k,m} \le 1 \qquad \forall i,j \tag{4}$$

$$Z_{i,j}^{k,m} \le x_k \qquad \forall i,j,k,m \tag{5}$$

$$Z_{i,j}^{k,m} \le x_m \qquad \forall i,j,k,m \tag{6}$$

$$Z_{i,j}^{k,m} \in \{0,1\} \tag{7}$$

$$x_k \in \{0,1\} \tag{8}$$

$$\text{Constraints for } \mathfrak{B} \tag{9}$$

$$\text{Constraints for } \mathfrak{C} \tag{10}$$

(1) is the objective functions, which try to maximize total constructor (government) incomes and minimize the public costs respectively, where $\mathfrak{B}$ is the earned benefit of establishing the hub system through tax/tolls. Technically $\mathfrak{B}$ is the counterpart of net present value (NPV) for a project with an infinite life cycle. In finance, NPV accounts for the time value of money and applies to a series of cash flows occurring at different times. The present value of a cash flow depends on the interval of time between now and the cash flow. It also depends on the discount rate. NPV calculations and tables for different periods and rate of return are shown in APPENDIX A.

As discussed earlier, $\mathfrak{B} - \mathfrak{C}$ cannot be used as an objective function, because $\mathfrak{B}$ is what constructors earn and in many cases, if it is not economically feasible, they don't build the system at all. On the other hand, $\mathfrak{C}$ is what customers pay to use the system, so a system which is attractive for both sides should be designed. Constraint (2) ensures that exactly P hubs are established. Constraint (3) considers the limit of the available budget. (4) Stipulates that each origin-destination (i,j) utmost could be allocated to only one pair of hubs (k,m). Surly this pair of hubs can be referred to just one hub because k and m could be the same. Constraints (5) and (6) assure that demand from origin i to destination j can be allocated to hub pair (k,m) if and only if both k and m are selected as the hubs. (7) and (8) define the decision variables as zero-one variables. (9) and (10) are needed constraints and calculations for constructors' benefits and users' costs respectively. These constraints may be varying for each policy. Below show these constraints for the first policy.

$$\frac{\sum_i \sum_j \sum_k \sum_m Z_{i,j}^{k,m}.(T_{k,m}.h_{i,j} - HC_{k,m})}{RR} - \sum_k x_k.hb_k \ \ge \mathfrak{B} \tag{11}$$

$$\sum_i \sum_j \sum_k \sum_m Z_{i,j}^{k,m}.h_{i,j}.C_{i,j}^{k,m} + \sum_i \sum_j \left(1 - \sum_k \sum_m Z_{i,j}^{k,m}\right) h_{i,j}.CC_{i,j} \le \mathfrak{C} \tag{12}$$

The (11) calculates the net present value (NPV) of the hub system with an infinite lifetime. Constraint (12) calculates the hub cost of users. $1 - \sum_k \sum_m Z_{i,j}^{k,m}$ is used to obviate the need of one extra variable and two extra constraints, which is used in the third model of (Kimms, 2005). This part has a vital role to mitigate the complexity of the problem.

2.1.2. Policy2: establishing an economic system

In this policy, in addition to public benefits, financial benefits are also a high priority. Constructor (/government) wants to reach the breakeven point in the very first period (year) and takes part only in economically feasible projects. So constraint (11) transforms into (13):

$$\frac{\sum_i \sum_j \sum_k \sum_m Z_{i,j}^{k,m}.(T_{k,m}.h_{i,j} - HC_{k,m})}{1 + RR} - \sum_k x_k.hb_k \ \ge \mathfrak{B} \tag{13}$$

To ensure that the established system is economically feasible, (14) is added to the model.

$$\mathfrak{B} \ \ge 0 \tag{14}$$

Note that RR^{-1} is changed to $(1 + RR)^{-1}$ because in (11) the life cycle of the project is infinite but in (13), the constructor wants to gain breakeven point within the first period so the life cycle should consider one.

2.1.3. Policy3: establishing an economic system within a makespan

The second policy suffers some weaknesses, in real-world cases; it is pretty improbable to achieve the breakeven point in the very first period (year). This case usually happens only if the establishing costs are quite low. Therefore, in Policy3, the constructor wants to establish a transportation system that reaches the breakeven point at most within MS periods. The modeling of this policy is the same as Policy(2) except (13) is turned into (15). Furthermore (16) and (17) are added to it.

$$\left(\sum_i \sum_j \sum_k \sum_m Z_{i,j}^{k,m}.(T_{k,m}.h_{i,j} - HC_{k,m})\right) \times (P/A, \ RR, t) - \sum_k x_k.hb_k \ge \mathfrak{B} \tag{15}$$

$$t \ \le MS \tag{16}$$

$$t \ge 1 \tag{17}$$

(15) and (16) guarantee that within MS (makespan) periods, the established system will achieve the breakeven point and it will be economically feasible. $(P/A, RR, t)$ is equal payment series present worth factor (a table based economic coefficient that converts annual net value (A) to the net present value (P) for t periods and at the rate of return of RR ; see APPENDIX A and for a thorough discussion, interested readers

are referred to (PANNEERSELVAM, 2013)). Constraints (16) and (17) are used to ensure that the number of periods doesn't exceed the maximum acceptable periods and also it is not less than one period. Because finding a suitable *MS* is a hard task for the decision-maker, the model can be extended by adding a new objective function as $\min MS$, to find the minimum *MS* internally. Then it will be compared with that one considered by the decision-maker. It's worth mentioning that *MS* those who are bigger than 10 years need some corrections and modification in demands, tax rates and other factors which are changed through time.

In APPENDIX B a numerical example is presented in which discount factors are calculated for each pair of nodes based on different factors such as the negative effect of distances. Further dissections on economies of scales are presented in section 3.

3. Economies of scale

The discount factor is one of the main drivers of installing hub systems, even so it did not gain enough consideration in previous studies. In classical hub location models, the hub-to-hub arcs are typically discounted by a fixed discount factor α, such that $0 \leq \alpha \leq 1$ (Alumur & Kara, 2008). However, some studies offered linear (Lüer-Villagra, Eiselt, & Marianov, 2019) or nonlinear (Alkaabneh, Diabat, & Elhedhli, 2019) formulations for discount factors. As stated earlier, discount factors are just calculated for the flow of traversing between hub nodes (hub-to-hub). However, some studies considered the hub-to-destination discount factor (Correia, Nickel, & Saldanha-da-Gama, 2018; Kimms, 2005). By the way, in real-world cases, the discount factors depend on many factors like the distance between each pair of nodes, the quality of paths, the type and capacity of the vehicle, the amount and type of commodity. In many cases, the relation of these factors with the discount factor is very vague. For example, in cases like the fragile goods shipping, nuclear wastes or any other hazardous cargo, the distance and amount may affect discount factors negatively. Surprisingly, in the literature, these negative effects are not studied and almost in all of them the distance and amount of cargo have a positive effect on discount factors and by increasing amounts, discounts are increased e.g. (Alkaabneh et al., 2019; Wagner, 2004). For all we know, no previous studies have been focused on the negative effects of distance and cargo amount and the effect of paths' quality and holding costs. These facts are the main motivation of developing new formulas for discount factors and taxes, based on items like distance, amount, cargo type, path's quality, fuel consumption, vehicle type, and more.

3.1. Transportation costs

To estimate discount and tax factors which are the most vital factors of establishing a new transportation system for constructors, enough knowledge about associated costs and benefits of the system is needed. As it is mentioned in the earlier section, many researchers used the amount and distance to estimate discount factors. However, the origin of costs is widely more than these factors, and ignoring them may cause fetal mistakes in estimations.

Some researches specifically focused on transportation costs, e.g. (Barnes & Langworthy, 2003; Indian roads congress, 2009; CEDEX, 2010; Ko, Lautala, Fan, & Shonnard, 2019), etc. (Barnes & Langworthy, 2003) presented their report about per-mile costs of operating automobiles and trucks. Their report breaks costs into five major components, including fuel, maintenance (excluding tiers), tiers, repairs, and deprecation. All results are presented in per-mile costs for distinctive classes of vehicles and different driving conditions e.g. pavements, and road type. Some adjustments are also included to cover specific situations that may occur. A group of researchers worked on research entitled socioeconomic and financial evaluation of transport projects. The results of this research gathered together as (CEDEX, 2010). It considered maintenance costs, operating costs (e.g. costs related to vehicles or assets, costs related to utilization time and costs related to distance traveled), and investment costs (e.g. planning costs, acquisition and land preparation costs, and construction costs) as the main costs of transportation projects. They also considered microeconomic variables like elasticities, maintenance costs, taxes, fuel prices, etc. In this report, the main factors of distance costs are fuel, tires, maintenance, and repairs. Indian roads congress (Indian roads congress, 2009) released a manual on economic evaluation of highway projects in India. They assert that roadway elements (e.g. pavement width, pavement type, vertical profile, etc.), vehicle factors (e.g. type, age, make, engine horse-power and power-weight ratio), and traffic considerations (e.g. traffic volume, traffic composition, speed, congestion) are most important parts affecting user costs. It also considers some benefits from highway improvements such as road user benefits (e.g. vehicle operating savings, savings in maintenance costs, etc.) and social benefits. All of these three research try to relate costs to distance, but distance without the amount of carried cargo is not enough. As the weight of shipped cargo is raised, the costs are also raised. Therefore, in this study in addition to the mentioned factors, the relations of costs to distance and weight of cargo are studied.

Based on the literature, in the viewpoint of users, transportation cost factors include fuel costs, vehicle maintaining, and repair costs, cargo assurance and maintaining costs, labor and administrative costs and taxes. In the viewpoint of constructors, in addition to users' cost factors, transportation costs include planning and research costs, constructing new paths, maintaining paths, applying controlling and supervision systems, etc. In the following, first, these factors are examined and then new formulations are presented to calculate discount factors of each path and taxes.

3.1.1. Fuel consumption

One of the most important factors in transportation costs is the fuel consumption cost (*FC*). The fuel consumption by itself depends on the type of vehicle (v), fuel type (e.g. petrol, diesel, CNG) and the unit cost ($FUC^{(v)}$), the fuel consumption rate per distance per weight ($FCR^{(v)}$), the type of cargo, the effect of the quality of paths ($PQF^{(v)}$) on fuel consumption, etc. Equation (18) shows the proposed formulation of fuel consumption for a given transportation system (SYS) e.g. the competing system and the hub system.

$$FC^{\mathrm{SYS}}{}_{k,m} = \left[FUC^{(v)}{}_{(\$)} \times FCR^{(v)}{}_{\left(\frac{1}{km \times kg}\right)} \times PQF^{(v)}{}_{k,m} \times d_{k,m_{(km)}} \right]_{\left(\frac{\$}{kg}\right)} \times \left[\sum_{\mathscr{K}} h_{i,j}{}^{(\mathscr{K})} \times UW^{(\mathscr{K})} \right]_{(kg)} \tag{18}$$

where $\mathscr{K}$ indicates the type of cargo and $UW^{(\mathscr{K})}$ is the unit weight of $\mathscr{K}$ th cargo. If there is only one type of cargo, Equation (18) can be written as below:

$$FC^{\mathrm{SYS}}{}_{k,m} = \left[FUC^{(v)}{}_{(\$)} \times FCR^{(v)}{}_{\left(\frac{1}{km \times kg}\right)} \times PQF^{(v)}{}_{k,m} \times d_{k,m_{(km)}} \times UW_{(kg)} \right]_{(\$)} \times h_{i,j} \tag{19}$$

It should be mentioned that based on (EUROPEAN ECONOMY, 2013) report, euro members tend to promote the use of diesel strongly through their relatively low tax rates in one hand, and on the other hand the cost of diesel is always lower than the petrol, so it could be concluded that day by day using diesel and vehicles which use diesel as fuel become more rational.

3.1.2. Vehicle maintenance and repair costs

The other important factor of transportation costs is the vehicle maintenance and repair cost (VMC). These costs include the costs of tiers, grease, etc. this type of cost can be related to the age of the vehicle, the quality of paths, distance, weight, etc. Equations (20) and (21) show the proposed formula for various and single cargo types respectively:

$$VMC^{SYS}{}_{k,m} = \left[PQCVM^{(v)}{}_{k,m} \times VAC^{(v)} \times VMR_{\left(\frac{\$}{km\times kg}\right)} \times d_{k,m_{(km)}}\right]_{\left(\frac{\$}{kg}\right)} \times \left[\sum_{\mathscr{K}} h_{i,j}{}^{(\mathscr{K})} \times UW^{(\mathscr{K})}\right]_{(kg)} \tag{20}$$

$$VMC^{SYS}{}_{k,m} = \left[PQCVM^{(v)}{}_{k,m} \times VAC^{(v)} \times VMR_{\left(\frac{\$}{km\times kg}\right)} \times d_{k,m_{(km)}} \times UW_{(kg)}\right]_{(\$)} \times h_{i,j} \tag{21}$$

where VMR is the cost of repair and maintenance per distance per weight at the factory condition (normal condition), $PQCVM^{(v)}$ and $VAC^{(v)}$ are coefficients of paths' quality and vehicle age on maintaining costs respectively which are used to modify normal conditions. As you see, the tire cost, repair costs and maintaining costs are combined, because the natures of all these factors are the same, and each class of the vehicle can be related to distance, age and weight of cargo.

3.1.3. Cargo assurance and holding costs

Broken cargo has no good for anyone, so the transportation system must deliver them whole and sound. Based on DaCoTA (DaCoTA, 2012), vehicle design and road safety are two important factors in freight safety. The other factors are the type of cargo and distance because each commodity based on its characteristics like the price, fragileness, hazardousness, etc., has some risks to be broken and as the distance becomes longer this risk is rising. The following equations show the proposed formula for cargo holding costs (CHC):

$$CHC^{SYS}{}_{k,m} = \left[PQCM^{(v)}{}_{k,m} \times SC^{(v)} \times d_{k,m_{(km)}}\right]_{(km)} \times \left[\sum_{\mathscr{K}} h_{i,j}{}^{(\mathscr{K})} \times CHR^{(\mathscr{K})}_{\left(\frac{\$}{km}\right)}\right] \tag{22}$$

$$CHC^{SYS}{}_{k,m} = \left[PQCM^{(v)}{}_{k,m} \times SC^{(v)} \times d_{k,m_{(km)}} \times CHR_{\left(\frac{\$}{km}\right)}\right]_{(\$)} \times h_{i,j} \tag{23}$$

where $PQCM^{(v)}$, $SC^{(v)}$ are the paths' quality coefficient and vehicle safety coefficient respectively. CHR is the cargo holding cost unit per distance. CHR has a higher value for hazardous, fragile and military cargo and lower value for typical freight.

3.1.4. Taxes and tolls

The other transportation cost origin is the tax. Governments and constructors typically use this means to mitigate and compensate for construction costs. Construction costs include planning costs, path constructing costs, path quality costs, controlling and other equipment costs, vehicle costs, labor costs, land costs, etc. as you may see, taxes have dual nature in transportation costs, in the viewpoint of customers, taxes are costs but at the same time they are benefits of constructors.

Each region and country calculating taxes and tolls have its own regulations such as (COMMISSION, 1997), but taxes always can be calculated based on the amount, the weight, the type of cargo, and the distance. For example, taxes on hazardous cargo are way more than typical cargo, and the longer distance causes higher taxes. The other benefits that constructors may gain are due to advertising contracts. These types of benefits can be considerable, but for the sake of brevity, they are not included in this study.

$$T^{SYS}{}_{k,m} = \left[d_{k,m_{(km)}}\right] \times \sum_{\mathscr{K}} TR_{\left(\frac{\$}{km\times kg}\right)}{}^{(\mathscr{K})} \times h_{i,j}{}^{(\mathscr{K})} \tag{24}$$

$$T^{SYS}{}_{k,m} = \left[d_{k,m_{(km)}} \times TR_{\left(\frac{\$}{km\times kg}\right)}\right] \times h_{i,j} \tag{25}$$

In the viewpoint of constructors, among these factors, planning costs, buying vehicles, equipment, and land costs can be considered as the fixed cost, the other one though can be considered as variable costs. For example, $HC_{k,m}$ can be calculated as below:

$$HC_{k,m} = d_{k,m} \times PMC \tag{26}$$

where PMC is the path maintaining cost per distance.

3.2. Discount factors

After a brief discussion on the origins of costs in transportation, more accurate economies of scale are presented below:

$$\alpha_{k,m} = \frac{TCH_{k,m}}{TCC_{k,m}} = \frac{FC^{hubSys}{}_{k,m} + VMC^{hubSys}{}_{k,m} + CHC^{hubSys}{}_{k,m} + T^{hubSys}{}_{k,m}}{FC^{CurrentSys}{}_{k,m} + VMC^{CurrentSys}{}_{k,m} + CHC^{CurrentSys}{}_{k,m} + T^{CurrentSys}{}_{k,m}} \tag{27}$$

where TCC and TCH are total costs of the current (/available) system and total costs of the hub system per weight per distance respectively. Note that calculation of χ and σ is the same as α with subtle modifications.

4. Proposed multi-objective algorithms

Even though integer programming optimization approaches are applied to solve small hub problems, larger instances of HLPs need to be solved by heuristic procedures or meta-heuristic procedures (Zanjirani Farahani, Hekmatfar, Boloori Arabani, & Nikbakhsh, 2013). Among these procedures, the genetic algorithm (henceforth referred to as GA) has been successfully used in solving many optimization problems (LuerVillagra & Marianov, 2013; Pham & Karaboga, 2000). In addition to well-known NSGAII (Deb, Pratap, Agarwal, & Meyarivan, 2002) that is widely used to solve HLPs, especially p-hub median problems (Ebrahimi Zade, Sadegheih, & Lotfi, 2014; Ghezavati & Hosseinifar, 2018), an improved version of an imperialist competitive algorithm is developed and applied to solve the problem. Although ICA has been used in some similar studies (Tavakkoli-Moghaddam, Gholipour-Kanani, & Shahramifar, 2013) but so far it has never been applied for multiple allocation p-hub median location problems. In the following subsections, a summary of some basic definitions and steps for each method and the way that they are improved and implemented is provided.

It worth mentioning that, because the problems are bi objective models, so it is the best to use multi-objective algorithms instead of changing the problem in single objective (for example weights cannot be used, because for the constructor, $\mathfrak{B}$ has a weight of 1 and C has a weight of zero in contrast for users it is opposite completely. Additionally, the lexicographic method cannot be used for the same reason. Each party tries to select the best option for him/herself.) and lose good options due to simplifications.

4.1. Non dominated sorting genetic algorithm (NSGAII)

GAs are population-based meta-heuristics which start with an initial population of solutions and enhance them by using some operators such as crossover and mutation. In the literature, some studies such as (Damgacioglu H., Dinler, Ozdemirel, & lygun, 2014; Damgacioglu H., Dinler, Ozdemirel, & Iyigun, 2014; Bashiri, Mirzaei, & Randall, 2013) used GA to solve HLPs. However classic GAs are used to solve only single objective, in the field of multi-objective optimization (Deb et al., 2002) proposed a new procedure based on GAs under title NSGAII. NSGAII is used to approximate a set of optimal solutions largely known as Pareto optimal solutions. Solutions in this set cannot be said to be better than the other. This approach is used to solve different variations of HLPs. For instance (Ebrahimi Zade et al., 2014) modified the NSGA-II by developing a dynamic immigration operator to solve a nonlinear multi-objective formulations for single and multiple allocation hub maximal covering problems as well as the linearized versions and gain better solutions and performance in compare to the traditional NSGA-II, and in (Ghezavati & Hosseinifar, 2018) an NSGA-II is utilized to solve a bi-objective hub facility location problem and they manage to achieve superior results using multi-objective particle swarm optimizers. In this study, a classic NSGA-II and an improved version of it will be applied to solve the model. Fig. 1 shows the pseudo-code of the proposed algorithm and, the following subsections demonstrate how it is implemented for presented problems.

4.1.1. Representation scheme

In all presented problems, there are two sets of decision variables, 1) the location of hubs and 2) assigning none hub nodes to these hub facilities. So two types of matrices are used to represent each of the mentioned variables. Type one as you see in Fig. 2 is facility location or hubs matrix. The algorithm randomly selects feasible nodes as hubs. Note that selected nodes must satisfy the number of hubs (constraint (3)) facilities and budget limitation (constraint (4)).

The second type matrix is built regarding the first type matrix. For each pair of origin/destination, two hubs are randomly selected as the first hub (FH) and the second hub (SH) and obviously they can be the same. Fig. 3 represents this type of matrix.

Then to complete the second type matrix, hub costs are compared with the best competitive system. If the competitive system has a lower cost than the hub system for one pair of origin/destination the customer uses it and vice versa. Fig. 4 shows this decision matrix. (See APPENDIX C for numerical example).

Note that the mask matrix is generally asymmetric because: 1) users in each side (of origin/destination) are different and ways to connect (roads and available systems) origins to destinations are not the same as destinations to origins. And 3) the type and amount of the cargo are not the same between origin/destination and vice versa. For instance, let's assume that we have only two nodes (A and B). Users on A side want to send cargo of type one and with the amount of X. He can use the established system or based on the condition maybe he uses the other systems. And his decision is completely independent (user B may not make the same decision because the available system may offer a better option than the established one on B to A than A to B for his cargo).

Also note that, the mask matrix is generated directly after proposing (generating) a solution matrix and it is used to check and calculate $CC_{i,j}$. If the proposed system offers a better cost than the rival ones, users choose that and the associate element of the matrix becomes 1 otherwise it becomes 0.

```
Procedure: The proposed Genetic algorithm optimization
Initialization mechanism
Repeat until the termination criterion is not met
{
    Parent selection mechanism
    Offspring generation mechanism
    Pareto front updating mechanism
}
```

Fig. 1. Proposed GA pseudo code.

hub_1	hub_2	...	hub_P

Fig. 2. Type one matrix (hubs matrix).

	1	2	...	N
1	-	$FH_{1,2}\rightarrow SH_{1,2}$	...	$FH_{1,N}\rightarrow SH_{1,N}$
2	$FH_{2,1}\|SH_{2,1}$	-	...	$FH_{2,N}\rightarrow SH_{2,N}$
⋮	...	...	-	...
N	$FH_{N,1}\rightarrow SH_{N,1}$	$FH_{N,2}\rightarrow SH_{N,2}$	...	-

Fig. 3. The second type matrix (assignment matrix).

	1	2	...	N
1	-	0	...	1
2	0	-	...	1
⋮	...	...	-	...
N	1	1	...	-

$$\begin{cases} 1 \; if \; min\left\{C_{i,j}^{k,m}, CC_{i,j}\right\} = C_{i,j}^{k,m} \\ 0 \; if \; min\left\{C_{i,j}^{k,m}, CC_{i,j}\right\} = CC_{i,j} \end{cases}$$

Fig. 4. Decision matrix.

4.1.2. Initialization mechanism

Since used algorithms are population-based approaches, they require a set of initial solutions. Frist, *PS*s individuals (see Table 2) are generated by randomly locate hubs and then assign each pair of origin/destination to exactly two hubs. Of course, these two hubs can be the same. To assign each origin to its destination, a simple heuristic is applied. The costs of sending cargo through, $i\rightarrow FH_{i,j}\rightarrow SH_{i,j}\rightarrow j$ and $i\rightarrow SH_{i,j}\rightarrow FH_{i,j}\rightarrow j$ are calculated and the path with lower cost is selected. When all associated costs of each pair of Origin/Destination are calculated, these values are compared with the best available alternative (in the viewpoint of customers), and the algorithm selects the optimum way to ship cargo from each pair of origin/destination. At last with regard to selected policy and constraints, overall fitness functions are calculated.

To be more specific, constraints (2) and (3) are taken care of in the type one matrix presented in Fig. 2 where only P nodes are selected randomly and if the total needed budget exceeds the available budget another set of P nodes are selected until the total needed budget meets available budget constraint. Constrains (4), (5) and (6) are taken care in the second type matrix depicted in Fig. 3 where for each pair of origin destination two hubs ($FH_{i,j}$, $SH_{i,j}$) randomly selected from the hub nodes (the type one matrix) this guarantees that a node is selected as hub node only if it was selected as a hub node, and because the algorithm does this for each pair of origin/destination only once, constraint (4) is satisfied. Constraints (9) and (10) are calculated constraints based on the type one and the type two matrices, and calculate fitness functions for each policy. Note that we used these two constraints as equations and multiplied (9) in -1 to find $-\mathfrak{C}$ ($-\mathfrak{C}$ is used in order to consider both fitness

Table 1
The strategy of generating used parameters.

The hub system	The competing System
Path maintaining cost (*PMC*) [dollars/Kilometer]	
U(50, 200)*16* U(21,30)	0
Tax/toll coefficient of using each kilometer of a path (*TR*) [cents/kilometer.kilogram]	
U(0.07,0.17)	0
Fuel cost based on used vehicle ($FUC^{(v)}$)[dollar/gallon]	
1.58	1.65
Fuel consumption based on used vehicle rate ($FCR^{(v)}$) [liter/kilometer.kilogram]	
37.9 lit/(100 km*50000 kg)	13.2 lit/(100 km*4100 kg)
The effect of quality of each path on vehicle fuel consumption (${PQF^{(v)}}_{k,m}$)	
U(0.7, 1.1)	$U(1.3,\ 1.8) \times {PQF^{(hub\ v)}}_{k,m}$
The unit weight of each cargo (*UW*) [kilogram]	
1	1
The effect of paths quality on maintaining costs (${PQCVM^{(v)}}_{k,m}$)	
U(0.7, 1.1)	$U(1.3,\ 1.8)\ \times {PQCVM^{(hub\ v)}}_{k,m}$
The effect of vehicle age on maintaining costs ($VAC^{(v)}$)	
U(1, 1.5)	U(1, 1.5)
The cost of repair and maintenance in factory condition (*VMR*) [dollar/kilometer. kilogram]	
U(50000, 30000)$*U(3,6)/ (21000 km*50000 kg)	U(500, 2000)$*U(10,15)/(10500 km*4100 kg)
The effect of path on cargo safety (${PQCM^{(v)}}_{k,m}$)	
U(1, 1.7)	U(1, 1.7)
The effect of vehicle type on cargo safety ($SC^{(v)}$)	
U(0.5,1)	U(0.8,1)
cargo holding cost unit per distance (*CHR*) [dollar/kilometer]	
0.2 $/100 km	0.5$/100 km
Initial budget (*HB*)	
U(0.1,0.3)*summation of all fixed costs	0

Table 2
Factors and levels.

Algorithm	Factors	symbol	levels		
Proposed GA	Population size	PS	50	100	150
	Probability of first type crossover operator	CW	0.7	0.8	0.9
	Probability of second type crossover operator	CS	0.7	0.8	0.9
	Probability of strong mutation operator	MS	0.1	0.2	0.3
	Probability of weak mutation operator	MW	0.1	0.3	0.5
	Probability of neighborhood search operator	*PLS*	0.8	0.85	0.9
	Number of iterations	NI	150	250	350
Proposed ICA	Number of countries	Ncnt	50	100	150
	Number of imperialists	Nimp	5%	10%	15%
	Assimilation coefficient	Beta	50%	60%	70%
	Deviation to assimilation	Teta	10%	20%	30%
	Revolution coefficient	RC	10%	15%	20%
	Colonies' power coefficient	Xi	0.05	0.10	0.15
	Number of iterations	NI	150	250	350

functions as maximization). Then in each iteration we tried to optimize $-\mathfrak{C}$ and $\mathfrak{B}$. Also note that to find which system will be used by users (the available one or the established one) we used decision matrix presented in Fig. 4.

4.1.3. Parent selection mechanism

In this study, binary tournament selection is used as the parent selection method. In this method, two random parents from all available parents are selected. Then they compared based on their ranks. A parent with a better rank (lower rank) is selected. If both selected parents have the same rank, they will be compared based on crowding distance value and the parent with greater crowding distance value will be selected. The concepts of rank and crowding distance described in (Deb et al., 2002) and for the sake of brevity, we do not duplicate them here.

4.1.4. Offspring generation mechanism

New individuals or offspring are generated by the means of two types of crossover and two types of mutation operators. Furthermore, a parallel neighborhood search is employed to gain better solutions from an existing one. The following subsections explain these operators and mechanisms.

4.1.4.1. Crossover. This paper considers two types of crossovers. In the first type (*CW*), a random zero-one mask matrix with size $N \times N$ is generated. Then arrays inside two selected parents are combined with regard to arrays of the mask matrix. For instance, if an array of the mask matrix is equal to one, the path will be selected from the first parent, and if it is equal to zero, the path will be selected from the second parent. The second type of crossover operator (*CS*) applies to the hub facility location matrix (type one). A zero-one mask matrix with size $1 \times P$ is used. Then like first type crossover parents are combined. These changes will spread through the assignment matrix.

4.1.4.2. Mutation. As the crossover operation, the mutation operator also has two types namely weak and strong mutation. In strong type (*MS*), the mutation occurs within the hub matrix (or type one matrix). This operator selects one element of the hub matrix and changes it with a new nod. Then this change spreads through the assignment matrix (or second type matrix) and changes all arrays that contain the removed hubs. Weak type mutation (*MW*), only affects the second type matrix. It selects one random array and changes associated *FH* or/and *SH* with existing hubs in the first matrix.

4.1.4.3. Parallel neighborhood search. As a hybrid strategy to further improve the performance, for any individual, a modified variable neighborhood search (*PLS*) (Ilić, Urošević, Brimberg, & Mladenović, 2010) is applied. Two structures of neighborhoods are utilized, which are referred to as allocate local search, and locate local search. To target customers' needs, an allocation procedure (Fig. 5) is applied. It randomly selects an element of the second matrix and changes associated hubs and if a better solution is gained, it will be replaced with the current one. Note that to target more demands, a roulette wheel selection procedure is used, to increase the chances of selecting a pair of origin/destination with higher $h_{i,j}$.

Primarily, locate local search (Fig. 6) is applied to consider constructors' satisfactions. It selects an element within the hub matrix randomly and a non-hub node using roulette wheel selection (nods with

```
Procedure: allocate local search
loop = 0
do {
    select an element of assignment matrix using roulette wheel sel
    randomly select two hubs from hub matrix
    generate new solution
    if (ℭnew < ℭcurrent) {
        current solution = new solution}
    loop ++
} while (Loop < N/2)
```

Fig. 5. Allocate local search procedure.

lower fixed required budget hb_k have more chance to select), then the hub node changes with non-hub one, and the effect spread through assign matrix. If this change has a positive effect on any of the objective functions, it will be accepted. For the sake of time consumption issues, only if the allocated procedure didn't gain a better solution, locate procedure will be applied.

4.1.5. Pareto front updating mechanism

To update the Pareto front, it is needed to compare new solutions with existing ones in Pareto front, so a fast non-dominated sorting and crowding distance procedures are applied, which introduced by (Deb et al., 2002). With these two procedures the rank and the crowding distance of each solution are calculated. Solutions with a minimum ranking number dominate the other ones and form the first Pareto front, then the remaining form the other fronts using the ranking number. The crowding distance value assures that when solutions with higher diversity are chosen. With regard to these two procedures, non-dominated solutions will be added to the Pareto front set and also dominated one will be removed from it.

4.2. Imperialist competitive algorithm (ICA)

The Imperialist Competitive Algorithm (ICA) was introduced by (Atashpaz-Gargari & Lucas, 2007) and in brief, is a new population-based algorithm in evolutionary computation that is inspired by a socio-political process of imperialistic competition in real-world. ICA starts with an initial population, named countries (counterpart of chromosomes in GA) then some of the best countries are chosen to be the “imperialists” and the rest are called “colonies”. These colonies are distributed among the imperialists based on their power (the more powerful an imperialist is, the more it occupies the colonies). Each of imperialists and their associated colonies, form an “empire” and the power of each empire is defined by the sum of the power of the imperialist and a percentage of the mean power of its colonies. Then the imperialistic competition begins among all empires, any empire which can't succeed to increase its power in these competitions is eliminated and the successful ones gradually increase their power by occupying colonies of other empires. Each imperialist tries to assimilate its colonies and make them a part of itself, this process is called assimilation. In this process, the characteristic of colonies tends to be close to the imperialist. Moreover, the power of some countries might change suddenly due to internal revolutions. ICA simulates these revolutions in colonies by generating some countries and replacing them with available countries, or changing their characteristics randomly. At last in each empire, if a colony gains more power than its associated imperialist the position of them is swapped and the colony becomes the imperialist and vice versa. Fig. 7 shows the proposed ICA and the following subsections describe how it is implemented.

4.2.1. Initiate empires mechanism

To generate initial N_{cnt} countries, the exact procedure in subsection 4.1.2. is used. the representation scheme is also the same in all algorithms (see subsection 4.1.1.). To initiate empires, for each country, first non-domination rank and crowding distance are calculated by applying the same procedure that is used in NSGAII and sort the countries in descending order. Top N_{imp} of sorted countries are defined as imperialists and the rest N_{col} are considered as colonies ($N_{cnt} = N_{imp} + N_{col}$). Note that to sort population the first criterion is non domination rank, and if two countries have the same rank, the better country is the one with bigger crowding distance. so the power of each imperialist is calculated as below:

$$p_i = \frac{\max\{R_i\} + \varepsilon}{R_i} + \frac{1}{CDI_i} \tag{28}$$

where R_i is the non-domination rank of ith imperialist, CDI_i is the index (/position) of crowding distance after sorting it and ε is a small positive number like 0.5. Normalized power of each imperialist is defined as:

$$NP_i = \frac{p_i}{\sum_{j=1}^{N_{imp}} p_j} \tag{29}$$

The initial colonies are divided randomly among empires based on imperialists' normalized power. The initial number of colonies of the ith empire will be

$$NoC_i = round(p_i \times N_{col}) \tag{30}$$

4.2.2. Assimilation mechanism

After assigning colonies to empires, each imperialist starts to improve their colonies. In order to assimilate their characteristics with the associated imperialist, colonies partially move toward their imperialist, and this movement is shown in Fig. 8.

As Fig. 8 shows, the movement has two major parameters r and θ. r is the movement radius, that is used to move toward the imperialist and is calculated as $U(0, \beta \times d)$ where β is the assimilation coefficient and d is the distance or difference between imperialist and colony. θ is the deviation to the direction of movement and it is used to search different points around the imperialist. In our discrete problem, this deviation is applied to the algorithm using strong and weak types of mutations that are introduced in subsection 4.1.4.2.

4.2.3. Revolution mechanism

As it is mentioned before, In ICA, the revolution can be simulated by generating some countries and replacing them with available countries. This revolution only occurs among colonies. This mechanism prevents the algorithm for sticking to a local optimum. To revolutionize colonies, the parallel local search introduced in subsection 4.1.4.3. is applied.

```
Procedure: locate local search
loop = 0
do {
    randomly select an element of hub matrix
    select a new non hub nod using roulette wheel selection
    generate new solution
    if (𝔅_new > 𝔅_current or ℭ_new < ℭ_current) {
        current solution = new solution}
    loop ++
} while (Loop < N/2)
```

Fig. 6. Locate local search procedure.

```
Procedure: The proposed multi-objective ICA
Generate initial countries
Initiate empires mechanism
Repeat until termination criterion is not met {
        Assimilation mechanism
        Revolution mechanism
        Swap and Update mechanism
        Competition mechanism
        Pareto front updating mechanism
}
```

Fig. 7. Proposed ICA pseudo code.

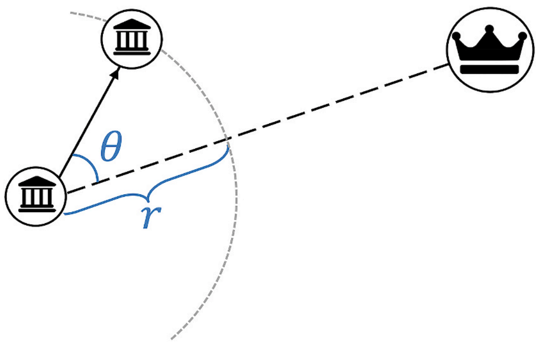


Fig. 8. Moving colony toward its imperialist.

4.2.4. Swap and update mechanism

After assimilation and revolution, a colony might reach a better position relevant to its imperialism. In this case, ICA replaces the imperialist position with the new powerful colony and the empire finds a new imperialist.

4.2.4. Competition mechanism

All empires try to take ownership of other empire's colonies. In ICA, the weakest colony of the weakest empire is selected and the competition between other empires starts to own this colony. All other empires have the opportunity to get the weakest colony, however, a more powerful empire has a higher change in this regard. In this study, first, the total power of each empire n is calculated as below:

$$TP_n = Power(imperialist_n) + \xi \times mean\{Power(colonies\ of\ empire_n)\} \tag{31}$$

where ξ is the coefficient of colonial power that affects the empire's power. Low ξ value leads the total cost of an empire to be closer to the imperialist cost. Moreover, an increase in this coefficient increases the impression of the mean colonies' power. After calculating the power of each empire, the roulette wheel selection is applied in order to select the winner empire. In these imperialistic competitions, empires with no colony will be eliminated and will be rolled as a colony.

5. Computational experience

In this study, a well-known database called AP is used. The problems are solved by the means of some meta-heuristic methods then compared with each other. Note that for the sake of brevity, the problem is solved only based on the first policy (See subsection 2.1. 1.). All algorithms are coded in Matlab 8.6.0 and executed on a PC with Intel(R) Xeon (R) CPU E5-1650 3.20 GHz CPU, and 32 GB of RAM on Windows 10 64bit operating system. To achieve the best possible results, it is needed to examine solutions based on quality and diversity. In the other words, in multi-objective optimization a good solution should have two important characteristics: 1) approximated Pareto front should converge to optimal Pareto front and 2) diversity maintaining of approximated Pareto front. To measure the quality and diversification of solution, hyper-volume I_H, additive epsilon $I_{\varepsilon+}$ indicators (Zitzler, Thiele, Laumanns, Fonseca, & Da Fonseca, 2003) and spacing metrics (Deb et al., 2002) are used respectively. The hyper volume index is defined as the volume of the objective space dominated by approximated Pareto front. In this study, the inclusion-exclusion algorithm proposed by (Wu & Azarm, 2000) is used. The ε-indicator gives the factor by which an approximation set is worse than another with respect to all objectives, or to be more precise: $I_\varepsilon(A,B)$ equals the minimum factor ε such that for any solution in B there is at least one solution in A that is not worse by a factor of ε in all objectives. In this study an additive epsilon indicator is used. A spacing metric is a measurement for the uniformity of the approximated Pareto front. Following formulation shows how to calculate this metric:

$$SP = \frac{\sum_{l=1}^{|PF|} |d_l - \overline{d}|}{(|PF| - 1)\overline{d}} \tag{32}$$

where $|PF|$ is the size of approximated Pareto front vector, d_l is the Euclidean distance between two consecutive solutions l and $l+1$ and $\overline{d}$ is the average of these distances.

5.1. Experimental data

Proposed algorithms have been tested on a standard test set known as AP (Australian Post), which is shared by (Ernst & Krishnamoorthy, 1996). In the literature discount factor values have been widely used: The collection cost factor χ, transfer cost factor α, and distribution cost factor σ are 3, 0.75 and 2, respectively. As it is mentioned in section 3 using fixed discount parameters is legitimate for real-world cases. So, AP test data is modified, and use presented formulations in section 3. Based on (Teck Sim, 2007) fixed costs considered to be random numbers between 20,000 and 200,000 and $RR = 10\%$. The strategy of generating other parameters is shown in Table 1. Note that, the data associated with vehicles are calculated based on technical specifications of two widely used vehicles in micro-shipments (the name of these two vehicles are removed due to branding issues) and the other ones are generated based on an established hub system in Iran (See APPENDIX B).

Also, note that we tried some multi objective algorithms like frog leaping algorithm, ant colony algorithm, bee colony algorithm and Imperialist Competitive Algorithm on a test set of problems (30 problems). As the ICA results were promising, we used it in current study in addition to improved NSGA-II.

5.2. Calibrations

Due to the significant effect on performance, setting and calibrating, the parameters of algorithms is an important job and multiple studies e.g. (Bernal, Castillo, Soria, & Valdez, 2020; 2017) used different methods to find the most optimized parameters for their algorithms. Because of a large number of parameters and factors, these calibrations require extensive experimentation; Therefore, the Taguchi method (Taguchi, 1986) is applied that uses fractional factorial experiments instead of full factorial one. In this approach, the response variable is converted to the signal-to-noise (S/N) ratio. As it is mentioned earlier hypervolume value is used as the solution quality indicator, and the reference point is considered to be the worst possible solution (minimum gained the benefit of the solutions and maximum gained cost of solution) so the bigger I_H indicates better quality. In the Taguchi approach for maximization problems, the following definition for S/N is used.

Due to the significant effect on performance, setting and calibrating, the parameters of a genetic algorithm is an important job. Because of a large number of parameters and factors, these calibrations require extensive experimentation; Therefore, the Taguchi method is applied that uses fractional factorial experiments instead of full factorial one. In this approach, the response variable is converted to the signal-to-noise (S/N) ratio. As it is mentioned earlier hyper volume value is used as the solution quality indicator, and the reference point is considered to be the worst possible solution (minimum gained benefit of the solutions and maximum gained cost of solution) so the bigger I_H indicates better quality. In the Taguchi approach for maximization problems, the following definition for S/N is used.

$$SN_l = -10 \log\left(\frac{1}{N_l}\sum_{u=1}^{N_l}\frac{1}{y_u^2}\right) \tag{33}$$

where y_u is the hypervolume indicator for a given experiment, and N_l is the number of trials for trial.

Each one of the proposed algorithms has seven control factors. Table 2 shows the considered levels of these factors. L_{18} is the most suitable orthogonal array for the proposed algorithms (see APPENDIX D). The mean S/N ratio of the proposed algorithm is shown in Table 2 and Fig. 7.

Table 3 also shows that the most important factors in the proposed genetic algorithm are the number of iterations (*NI*) and population size (*PS*). The best values of *PS*, *CW*, *CS*, *MS*, *MW*, *PLS* and *NI* are 100, 0.8, 0.8, 0.2, 0.3, 0.85 and 350 respectively. For the proposed ICA also the number of iteration and number of countries (*Ncnt*) are the most important factors. The best values of *Ncnt*, *Nimp*, *Beta*, *Teta*, *RC*, *Xi* and *NI* are 100, 10%, 60%, 20%, 15%, 0.10, and 350 respectively.

5.3. Results

This section compares the proposed genetic algorithm (PGA), the proposed imperialistic competition algorithm (PICA) and basic NSGAII, that are discussed in section 4. To compare these algorithms in the viewpoint of quality and diversity, we used hypervolume, epsilon indicators, and spacing metrics index respectively. In addition to that CPU time of these are also compared.

To bold and clarify the effect of using hub systems in real-world cases, the average use of the system by customers is also considered. Table 4 shows simulation results for the first policy (gaining break-even point within the first period). However, results of the third policy can be easily calculated using $(P/A, RR, t)$. As you see in the table in almost all cases the usage of the hub system is lower than 50 percent, which means designers and constructors of the hub system cannot achieve the entire transportation demand. And if constructors use classical hub models and assume that they can gain the whole demand, the designed system probably causes a high level of finance wastes.

Table 4 also shows, proposed ICA is 10% and 19% better than proposed GA and NSGAII in regard to hypervolume indicator, 2.63 and 3.41 times better than proposed GA and NSGAII in regard to epsilon indicator and it gives 10% and 17% better solutions in regard with spacing metric. It is also 60% and 15% less time consuming than proposed GA and NSGAII respectively. Table 4 reveals that in a competitive atmosphere, constructors can lose more than 50 percent of shipments, and their investments can be totally collapsed if the business models don't separate their costs and benefits from users'. To understand how constructors can take advantage of taxes and select the most profitable plan, we studied the tax versus usage and its effect on constructor's benefits. Fig. 9 (a) makes it obvious that, even if constructors want to follow the first policy and exempt all taxes, the usage still won't get to 100%, because in some cases rival systems can be less costly for users than the hub system. As it is shown in the illustrative example (APPENDIX B), these cases almost occur for close distances. The figure also shows if constructors want to follow two other policies, they can only increase taxes to a threshold. For example, in the simulated conditions taxes more than 52 cents per kilometer per kilogram will collapse the whole system and its usage will downgrade to zero. Fig. 9 (b) shows how taxes can affect constructors' benefits. It depicts that initially tax increases have a positive impact of the $\mathfrak{B}$ but after crossing a threshold, increasing taxes will damage constructors' benefits. In the simulation condition this threshold is about 29 cents on average. The ragged parts of the figure show how the established system can resist tax changes. For example, 33 cents of tax show a better performance than 32 cents, because users can tolerate both values the same, so this 1 extra cent can increase $\mathfrak{B}$ by almost 3%.

Except for taxes another important factor on objective functions is the is the fuel consumption cost (*FC*). Note that the fuel used in rival systems (micro-transportation) and the proposed systems are generally petrol and diesel respectively and diesel's cost grows fairly slower than the petrol and also governments usually pay more subsidies for diesel in order to support mass transportation. For example, In Iran the price of diesel is more than 60% lower than the price of gasoline and the growth rates of petrol and diesel in average are 18.5% and 10% respectively. This difference can help proposed transportations a lot if the third policy is selected. Because each year more users shift from rival systems to proposed systems, so in this case constructors can earn more benefits if they invest their money on the third policy.

The other important factor that affects the way of establishment is the available budget. If the project is started with a low initial budget (*HB*) lots of good choices will be lost and constructors should accept only low cost nodes as the hub nodes. However, by raising the initial budget the more profitable facilities can be established these facilities also can increase usage more than the other ones.

Fig. 10 shows how the initial budget (see Table 1) affects the usage of the proposed system. If a low budget is available initially, constructors get in trouble for finding feasible places and inevitably establish hub facilities in nodes that need low budgets. Generally, these facilities are outlying centers that cannot handle lots of shipments and users are not interested in using them, so the overall usage and constructors' benefits stay low. On the other hand, after investing enough money, the best available options can be taken and investing more won't help the system. It is worth mentioning that low budget makes a lot of solutions unfeasible and can increase needed CPU time of simulations drastically.

From the perspective of algorithms, and with regard to Table 4 proposed ICA has superiority in both quality and diversity of solutions. To examine the superiority of the proposed ICA, an ANOVA test is applied. Before using the ANOVA, all data are tested for normality by the means of Kolmogorov-Smirnov test and it shows that the data are normal, so, using the ANOVA test is justified. ANOVA results shown in Table 5 indicate that the superiority of the hybrid ICA is significant in the 95% confidence level for the mentioned indexes. To visualize Pareto front for each algorithm, three figures of three sizes of the problem is shown in Fig. 11. This figure shows that for small size instances, the quality of solution for all algorithms is almost the same but as the size of the problem grows higher, the proposed ICA gives much better solutions than the two others in less time.

6. Conclusion and further research

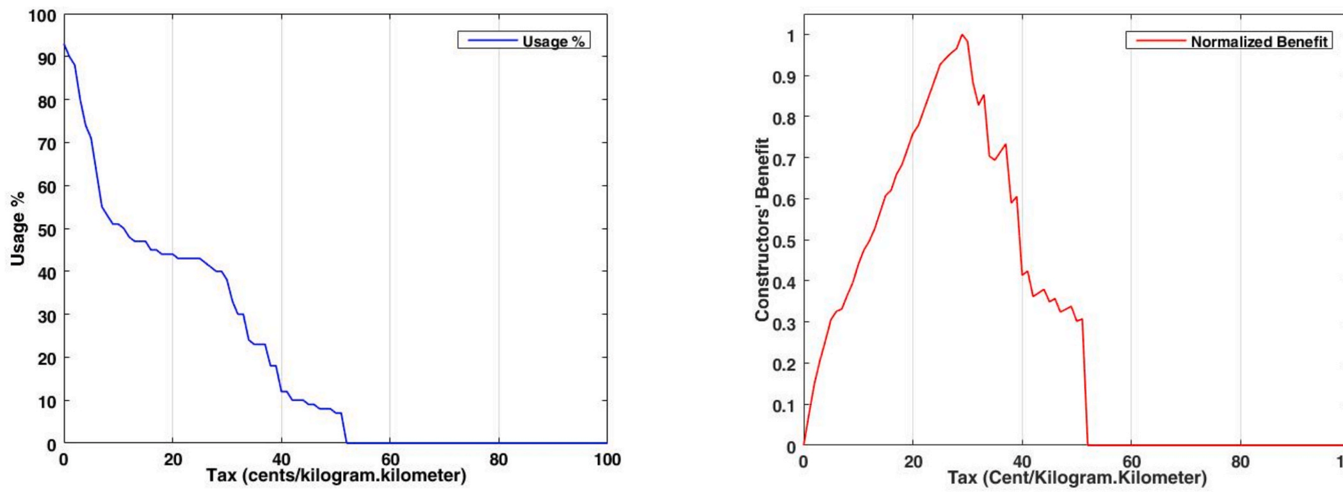
This paper studied the p hub median problem, concerning real-world assumptions and discussed different policies to install a new hub system in a competitive environment. Different shortages and drawbacks of previous models and assumptions are described. Particularly, the paper outlined boundaries between users' and constructors' costs and benefits and presented more robust economies of scales that can handle more real situations like shipping fragile goods, military equipment, and hazardous wastes wherein the distance and the amount of cargo may damage the conventional hub discounts alongside the normal case wherein the mentioned factors affect discounts positively. Also, it

Table 3
Best value of each factor.

Proposed GA	Factors	PS	CW	CS	MS	MW	*PLS*	NI
	Rank	2	7	3	5	6	4	1
	Best Value	100	0.8	0.8	0.2	0.3	0.85	350
Proposed ICA	Factors	Ncnt	Nimp	Beta	Teta	RC	Xi	NI
	Rank	2	6	4	7	3	5	1
	Best Value	100	10%	60%	20%	15%	0.10	350

Table 4
The summarized results obtained by the algorithms.

			Hyper-volume indicator			Additive epsilon indicator			Spacing metric			CPU Time			
Size	N	P	PICA	PGA	NSGAII	PICA	PGA	NSGAII	PICA	PGA	NSGAII	PICA	PGA	NSGAII	usage
Small	10 to 20	3 to 9	0.626	0.598	0.525	0.019	0.051	0.057	0.561	0.507	0.437	35.2417	63.3796	39.602	51.070
Medium	25 to 35	12 to 17	0.802	0.735	0.662	0.0205	0.070	0.087	0.578	0.494	0.407	115.522	206.463	131.704	50.075
Large	40 to 50	15 to 20	0.945	0.809	0.640	0.020	0.097	0.121	0.633	0.520	0.419	244.655	439.852	280.433	49.601



(a) Tax value vs usage percentage

(b) Tax value vs constructors' benefits

Fig. 9. Tax effect on usage and constructors' benefit.

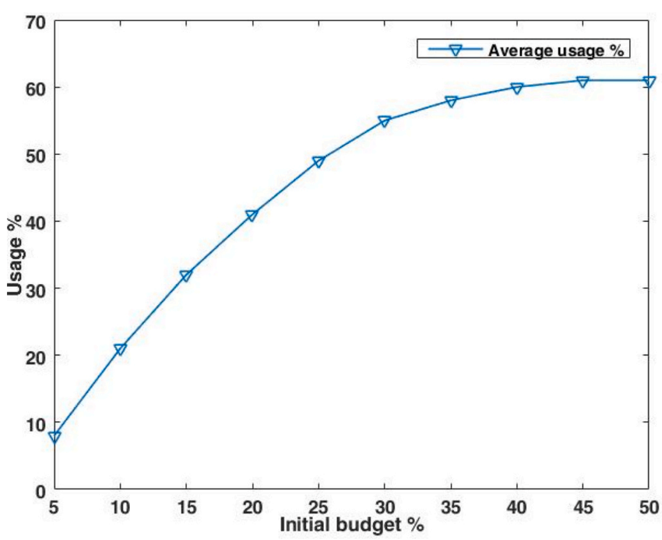


Fig. 10. The effect of the initial budget on system usage.

Table 5
Indicators mean value for algorithms and related p-value for each level.

Index	Source	DF	SS	MS	F	P
Hyper volume indicator	Algorithm	2	0.8735	0.4368	12.96	0.000
Additive epsilon indicator	Algorithm	2	0.13231	0.6616	51.53	0.000
Spacing metric	Algorithm	2	0.7260	0.3630	21.12	0.000

considered the availability of other competing transportation systems and discussed how these rival systems can degrade the share of the established system and earn benefits.

The proposed model can cover three possible policies wherein constructors either want to establish a hub system merely to facilitate the general transportation or participating in a profitable economic project and gaining break-even point within the first period or more. The model is solved using some enhanced meta-heuristic algorithms e.g. the non-dominated sorting genetic algorithm II (NSGA-II) and the imperialist competitive algorithm (ICA). These algorithms are finely tuned using the Taguchi method and compared with each other using indicators like hypervolume, additive epsilon, spacing metric, and CPU time and the results showed the outperformance of the hybrid imperialist competition algorithm. The simulations results unveiled that the rival systems can decrease shipment share of the established hub system more than 50% and the assumption of gaining the whole share, may cause drastic wastes of finance for constructors.

The research considered a fixed budget for establishing the hub system, but in reality, constructors can raise more funds to the project and also use the excess budget and invest it in the bank or use it to establish more hub facilities. Also in this model constructors should establish the whole system at once but in reality, a basic system can be built first and then it can grow gradually. Moreover, as it is mentioned in subsection 5.3 an important factor for the success of the proposed system is the tax value and if tax value can be modeled as an endogenous variable, it can help constructors' to choose the best plan. Accordingly, for further research, we recommend readers to consider the mentioned issues and develop models that can meet more real case conditions. Moreover, in this study, an unlimited hub and road capacity and the same weight and type of cargo are considered, but the model can be extended to consider different types of cargo with different weights and put some constraints on the capacity of the hubs and roads.

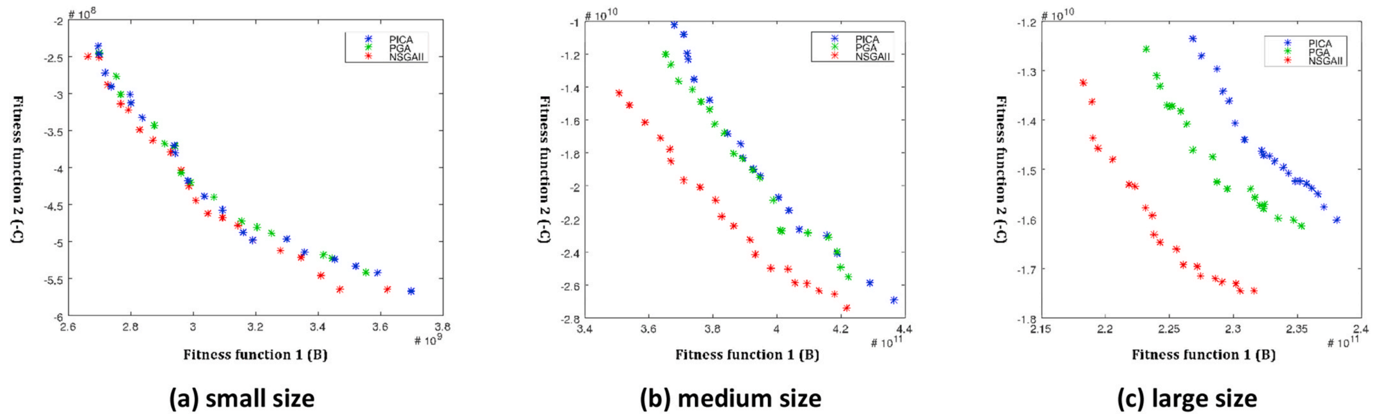


Fig. 11. Pareto front comparison for different sizes.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgements

Compliance with ethical standards.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Appendix A. Present worth of one dollar per period payable at end of each period

$$P = A\left[\frac{(1 + RR)^{t} - 1}{RR\,(1 + RR)^{t}}\right]$$

$$t = 1 : P = A\frac{1}{(1 + RR)}$$

$$t \rightarrow \infty : P = A\frac{1}{RR}$$

Years	6%	7%	8%	9%	10%	11%	12%	13%	14%
1	$0.943396	$0.934579	$0.925926	$0.917431	$0.909091	$0.900901	$0.892857	$0.884956	$0.877193
2	$1.833393	$1.808018	$1.783265	$1.759111	$1.735537	$1.712523	$1.690051	$1.668102	$1.646661
3	$2.673012	$2.624316	$2.577097	$2.531295	$2.486852	$2.443715	$2.401831	$2.361153	$2.321632
4	$3.465106	$3.387211	$3.312127	$3.239720	$3.169865	$3.102446	$3.037349	$2.974471	$2.913712
5	$4.212364	$4.100197	$3.992710	$3.889651	$3.790787	$3.695897	$3.604776	$3.517231	$3.433081
6	$4.917324	$4.766540	$4.622880	$4.485919	$4.355261	$4.230538	$4.111407	$3.997550	$3.888668
7	$5.582381	$5.389289	$5.206370	$5.032953	$4.868419	$4.712196	$4.563757	$4.422610	$4.288305
8	$6.209794	$5.971299	$5.746639	$5.534819	$5.334926	$5.146123	$4.967640	$4.798770	$4.638864
9	$6.801692	$6.515232	$6.246888	$5.995247	$5.759024	$5.537048	$5.328250	$5.131655	$4.946372
10	$7.360087	$7.023582	$6.710081	$6.417658	$6.144567	$5.889232	$5.650223	$5.426243	$5.216116
11	$7.886875	$7.498674	$7.138964	$6.805191	$6.495061	$6.206515	$5.937699	$5.686941	$5.452733
12	$8.383844	$7.942686	$7.536078	$7.160725	$6.813692	$6.492356	$6.194374	$5.917647	$5.660292
13	$8.852683	$8.357651	$7.903776	$7.486904	$7.103356	$6.749870	$6.423548	$6.121812	$5.842362
14	$9.294984	$8.745468	$8.244237	$7.786150	$7.366687	$6.981865	$6.628168	$6.302488	$6.002072
15	$9.712249	$9.107914	$8.559479	$8.060688	$7.606080	$7.190870	$6.810864	$6.462379	$6.142168
16	$10.105895	$9.446649	$8.851369	$8.312558	$7.823709	$7.379162	$6.973986	$6.603875	$6.265060
17	$10.477260	$9.763223	$9.121638	$8.543631	$8.021553	$7.548794	$7.119630	$6.729093	$6.372859
18	$10.827603	$10.059087	$9.371887	$8.755625	$8.201412	$7.701617	$7.249670	$6.839905	$6.467420
19	$11.158116	$10.335595	$9.603599	$8.950115	$8.364920	$7.839294	$7.365777	$6.937969	$6.550369
20	$11.469921	$10.594014	$9.818147	$9.128546	$8.513564	$7.963328	$7.469444	$7.024752	$6.623131
21	$11.764077	$10.835527	$10.016803	$9.292244	$8.648694	$8.075070	$7.562003	$7.101550	$6.686957
22	$12.041582	$11.061240	$10.200744	$9.442425	$8.771540	$8.175739	$7.644646	$7.169513	$6.742944
23	$12.303379	$11.272187	$10.371059	$9.580207	$8.883218	$8.266432	$7.718434	$7.229658	$6.792056
24	$12.550358	$11.469334	$10.528758	$9.706612	$8.984744	$8.348137	$7.784316	$7.282883	$6.835137
25	$12.783356	$11.653583	$10.674776	$9.822580	$9.077040	$8.421745	$7.843139	$7.329985	$6.872927

Appendix B. Illustrative example

In this section we use data of a real-world company as a numerical example. The company is a water desalination producer and importer which has 4 factories and 3 warehouses. Parts and production of this company are very fragile and expensive, so discount factors are affected negatively by distance (because longer distances may put cargo in higher risk of flaw. See section 3 for thorough discussion). A transportation company wants to facilitate shipments by establishing a new hub location system with regard to second or third policy and the fact that direct shipment in all paths are allowed. Off-course real data is not allowed to reveal but we use modified versions and one of its feasible solutions to disclose different aspects of the policies and associated mechanisms. Figure B1 shows the location of these 7 facilities. Maximum available budget is 453000 monetary units. RR is equal to 12% and there is no maintaining cost $HC_{k,m} = 0$. For each pair of nodes distances, shipments, discount factors and tax rates are shown in Table B1-Table B5. Please note that distance between each pair of origin-destination must be equal and cost of shipments must be dependent on distances, so we consider 1 monetary unit as the associated transportation cost per mile. There is no obligation for equality of shipments, α and σ between pairs of origin/destination (i,j) and (j,i), but for the sake of simplicity we considered them equal. As we mentioned earlier the more distance causes higher tax rate and lower discounts. Number of hub nodes is 2 and fixed cost of establishment of each hub is shown in Table B.6.

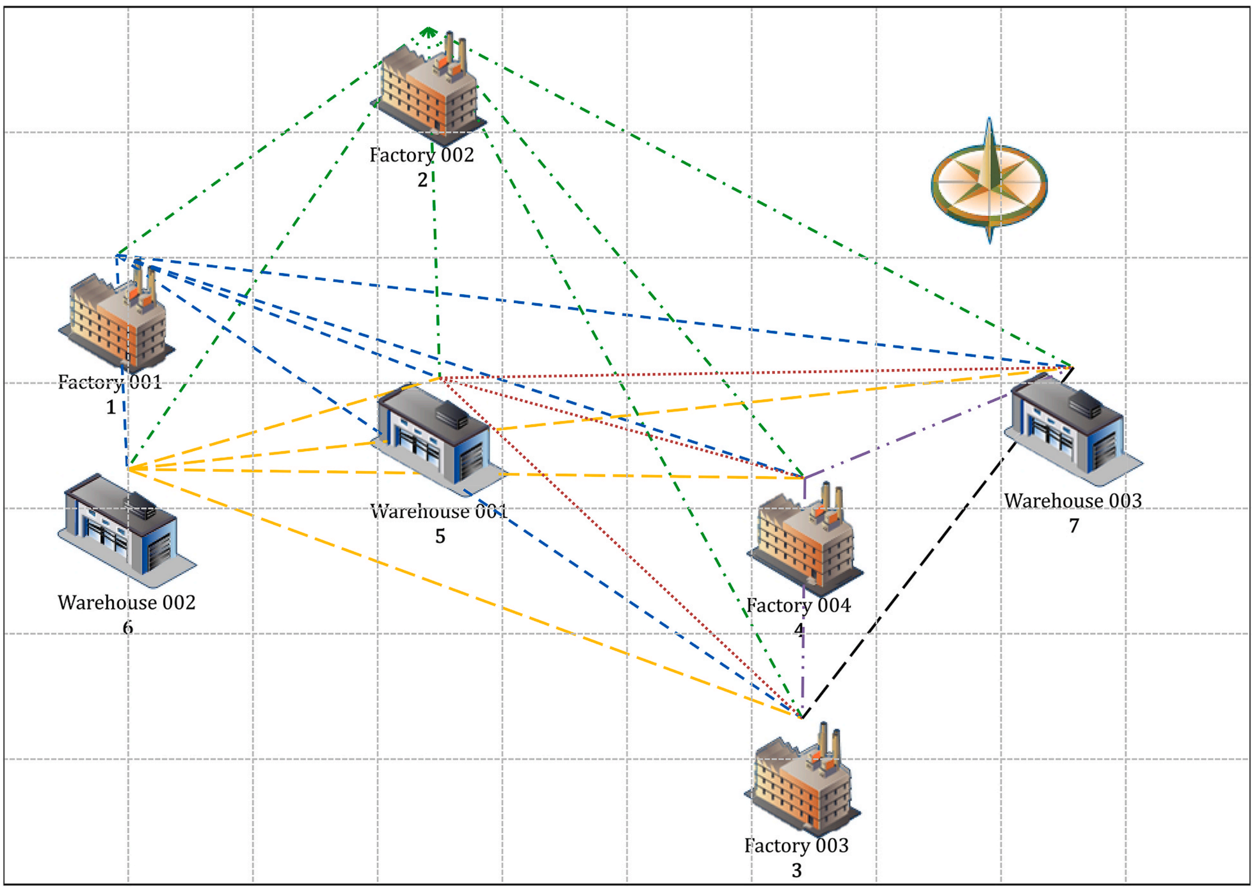


Fig. B.1. Locations of nodes.

Table B.1
distance between each pair of nodes *10 miles

	Factory1	Factory2	Factory3	Factory4	Warehouse1	Warehouse2	Warehouse3
Factory1	0.000	3.086	6.629	5.800	2.769	1.712	7.733
Factory2	3.086	0.000	6.274	4.691	2.791	4.272	5.846
Factory3	6.629	6.274	0.000	1.920	3.983	5.774	3.549
Factory4	5.800	4.691	1.920	0.000	3.037	5.440	2.332
Warehouse1	2.769	2.791	3.983	3.037	0.000	2.614	5.081
Warehouse2	1.712	4.272	5.774	5.440	2.614	0.000	7.633
Warehouse3	7.733	5.846	3.549	2.332	5.081	7.633	0.000

Table B.2
Total cargo must be delivered from origin to destination

	Factory1	Factory2	Factory3	Factory4	Warehouse1	Warehouse2	Warehouse3
Factory1	0	69625	38061	19361	95198	21413	73687
Factory2	69625	0	95406	78000	92821	14649	82351
Factory3	38061	95406	0	76145	50621	24263	12822
Factory4	19361	78000	76145	0	54036	19504	88037
Warehouse1	95198	92821	50621	54036	0	87089	23257
Warehouse2	21413	14649	24263	19504	87089	0	16525
Warehouse3	73687	82351	12822	88037	23257	16525	0

Table B.3
discount between each pair of hub nodes

	Factory1	Factory2	Factory3	Factory4	Warehouse1	Warehouse2	Warehouse3
Factory1	0.00%	71.23%	78.59%	77.67%	69.65%	60.18%	79.50%
Factory2	71.23%	0.00%	78.23%	75.94%	69.77%	75.05%	77.73%
Factory3	78.59%	78.23%	0.00%	62.87%	74.33%	77.64%	73.02%
Factory4	77.67%	75.94%	62.87%	0.00%	61.01%	77.19%	66.78%
Warehouse1	69.65%	69.77%	74.33%	61.01%	0.00%	68.74%	76.63%
Warehouse2	60.18%	75.05%	77.64%	77.19%	68.74%	0.00%	79.43%
Warehouse3	79.50%	77.73%	73.02%	66.78%	76.63%	79.43%	0.00%

Table B.4
discount between second hub and destination

	Factory1	Factory2	Factory3	Factory4	Warehouse1	Warehouse2	Warehouse3
Factory1	0.00%	83.32%	86.16%	85.30%	82.33%	73.60%	87.65%
Factory2	83.32%	0.00%	85.38%	83.98%	79.26%	83.10%	87.94%
Factory3	86.16%	85.38%	0.00%	75.94%	85.59%	87.19%	81.69%
Factory4	85.30%	83.98%	75.94%	0.00%	73.14%	84.58%	77.21%
Warehouse1	82.33%	79.26%	85.59%	73.14%	0.00%	78.84%	84.87%
Warehouse2	73.60%	83.10%	87.19%	84.58%	78.84%	0.00%	88.58%
Warehouse3	87.65%	87.94%	81.69%	77.21%	84.87%	88.58%	0.00%

Table B.5tax rate between each pair of hubs

	Factory1	Factory2	Factory3	Factory4	Warehouse1	Warehouse2	Warehouse3
Factory1	0.00%	1.54%	3.31%	2.90%	1.38%	0.86%	3.87%
Factory2	1.54%	0.00%	3.14%	2.35%	1.40%	2.14%	2.92%
Factory3	3.31%	3.14%	0.00%	0.96%	1.99%	2.89%	1.77%
Factory4	2.90%	2.35%	0.96%	0.00%	1.52%	2.72%	1.17%
Warehouse1	1.38%	1.40%	1.99%	1.52%	0.00%	1.31%	2.54%
Warehouse2	0.86%	2.14%	2.89%	2.72%	1.31%	0.00%	3.82%
Warehouse3	3.87%	2.92%	1.77%	1.17%	2.54%	3.82%	0.00%

Table B.6
fixed cost of hub establishment

Factory1	Factory2	Factory3	Factory4	Warehouse1	Warehouse2	Warehouse3
199000	195000	221000	193000	180000	245000	243000

First, we propose a feasible solution of the problem (approximated Pareto front solutions are presented in section 5.) considering common assumptions of hub systems, such as commodities must ship only via hub system and direct connections between the non-hub nodes are not allowed. Then, a solution in a more realistic way is proposed in which each node uses the best way to ship its commodities, such as a competing system (in this case direct system) or hub system. Associated results of these two approaches are illustrated in Figure B2 and Figure B3 respectively.

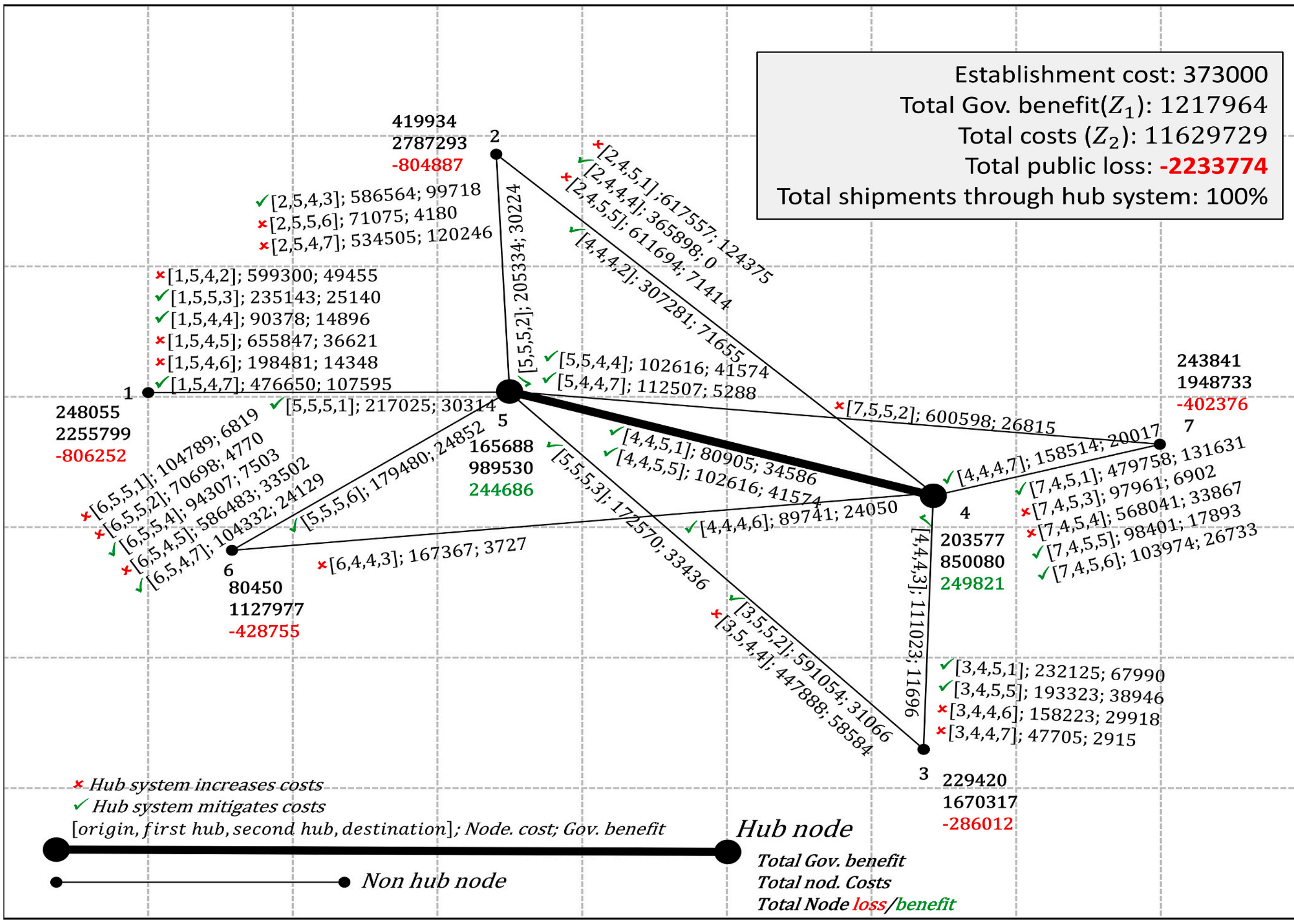


Fig. B.2. Classical hub transportation system without competitor.

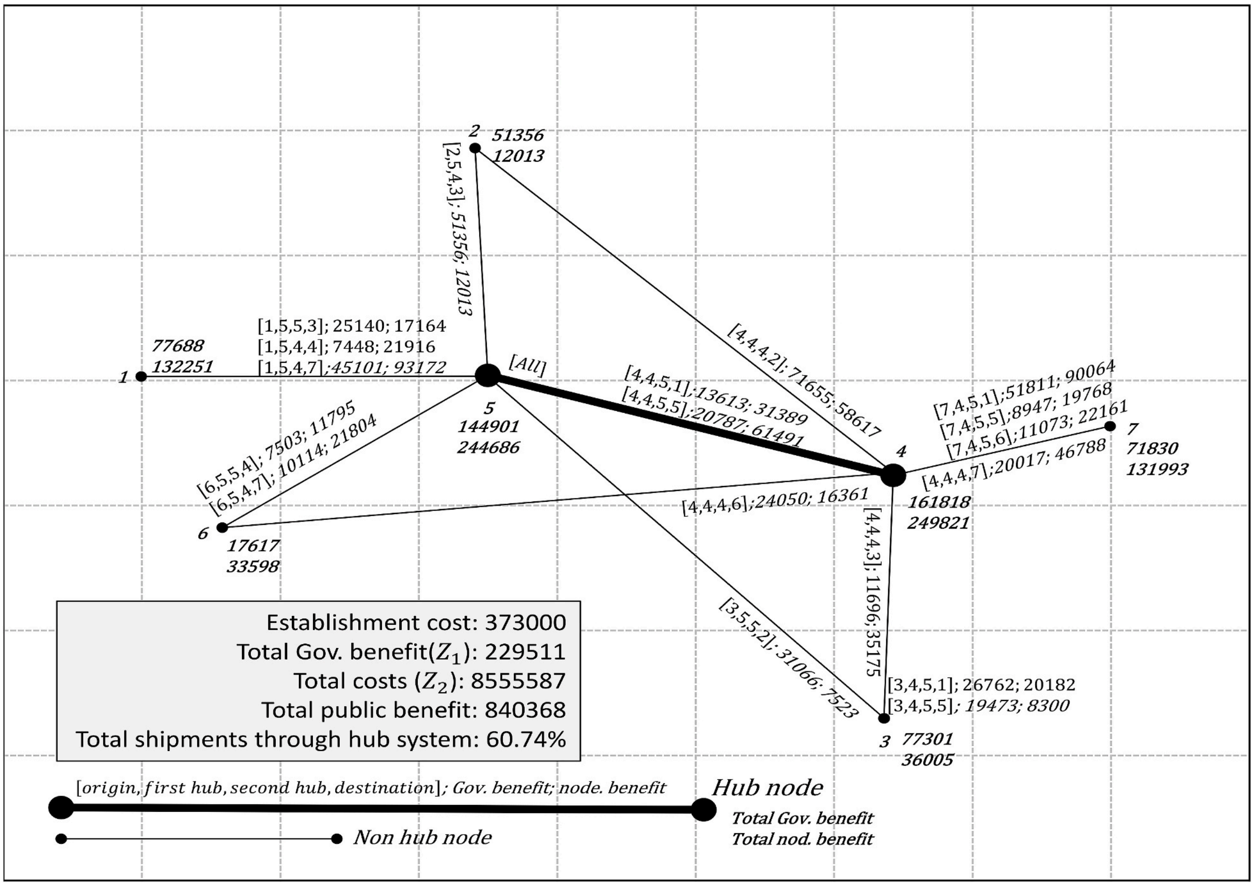


Fig. B.3. Hub system in competitive atmosphere.

As you see in Figure B3 hubs are nodes 4 and 5. total public costs in the proposed system is 8,555,587 that is 3,074,142 lower than classical hub system and 840,368 lower than competing systems, which means establishing a new hub system has benefits for both constructor and public. In one hand constructor can compensate its costs (establishment cost) in the very first period (second policy) and gain 229511. In other hand public cost is mitigated about 9% versus competing systems and 26% versus hub system separately. Note that in a realistic viewpoint only about 60% of commodities are shipped through the hub system and numerical expedients (see 5.3) show that in optimal solutions this usage percentage is even lower than that.

If nodes are forced to use established hub system (as it is in classic hub assumptions) constructor gains 1,217,964, that is 988,453 (4.31 times) more than proposed version (229511) but in real-world cases this is a very delusive result because public costs are increased 2,233,774 more than available system and 3,074,142 more than proposed system and undoubtedly nodes do not fully use established hub system. As you see in Figure B3 in many cases, the cost of available system is much lower than established system (e.g. shipment from origin 2 to destination 1 has cost of 214,863 through available system and 617,557 through hub system that is 1.87 times more than available system so node 2 sent all its commodities through competing system except commodities with destination 3).

Appendix C. Applying proposed Meta heuristic for illustrative example

This matrix for illustrative example is shown in Figure C1.

4	5

Fig. C.1. Hub matrix for illustrative example.

For the illustrative example assignment matrix is shown in Figure C2.

	1	2	3	4	5	6	7
1		5→4	5→5	5→4	5→4	5→4	5→4
2	4→5		5→5	4→4	4→5	5→5	5→4
3	**4→5***	5→5		5→4	4→5	4→4	4→4
4	4→5	4→4	4→4		4→5	4→4	4→4
5	5→5	5→5	5→5	5→4		5→5	4→4
6	5→5	5→5	4→4	5→5	5→4		5→4
7	4→5	5→5	4→5	4→5	4→5	4→5	

Fig. C.2. assignment matrix of illustrative example.
*For example cargo from origin 3 shipped through first hub 4, then second hub 5 and at last destination 1.

For the illustrative example this decision matrix is illustrated in Figure C3.

	1	2	3	4	5	6	7
1		0	1	1	0	0	1
2	0		1	1	0	0	0
3	1	1		0	1	0	0
4	1	1	1		1	1	1
5	1	1	1	1		1	1
6	0	0	0	1	0		1
7	1	0	0	0	1	1	

Fig. C.3. decision matrix for illustrative example.

Table C1 and Figure C4 show the results of both algorithms for illustrative example (see APPENDIX B).

Table C.1
Approximated Pareto of Illustrative example

Hubs	PGA		Hubs	NSGAII	
	Benefit	Cost	Hubs	Benefit	Cost
1, 7*	1919428	9026330	1, 7*	1898277	9049355
1, 2	1828315	8827137	1, 3	1876566	8914937
2, 4*	1491085	8666700	2, 3	1831172	8856956
4, 5	1374141	8473624	2, 4*	1410797	8672362

As you see in Table C1 both algorithms found 1, 7 and 2, 4 as hub locations, but the proposed algorithm allocates spokes to hub nodes better than NSGAII, so with the same hub nodes, better benefits and costs are gained. The superiority of the proposed algorithm is clear in Figure C4 because it gains better hyper volume (after standardization: 0.70 to 0.63) and it proposes more diverse solutions (0.24–0.71).

Table C.2
Total customer costs and total constructor costs during time for illustrative example

Plan	Hubs	TC	N = 1	N = 3	N = 5	N = 10	N = 20	N = ∞
1	1,7	9.03E+06	8.21E+06	2.24E+07	3.42E+07	5.55E+07	7.68E+07	9.03E+07
2	1,2	8.83E+06	8.02E+06	2.20E+07	3.35E+07	5.42E+07	7.52E+07	8.83E+07
3	2,4	8.67E+06	7.88E+06	2.16E+07	3.29E+07	5.33E+07	7.38E+07	8.67E+07
4*	4,5	8.47E+06	7.70E+06	2.11E+07	3.21E+07	5.21E+07	7.21E+07	

(continued on next page)

Table C.2 (*continued*)

Plan	Hubs	TC	N = 1	N = 3	N = 5	N = 10	N = 20	N = ∞
								8.47E+07

Plan	**Hubs**	**TB**	**Fixed Cost**	**N = 1**	**N = 3**	**N = 5**	**N = 10**	**N = 20**	**N = ∞**
1*	1,7	9.03E+06	4.42E+05	1.30E+06	4.33E+06	6.83E+06	1.14E+07	1.59E+07	1.88E+07
2	1,2	8.83E+06	3.94E+05	1.27E+06	4.15E+06	6.54E+06	1.08E+07	1.52E+07	1.79E+07
3	2,4	8.67E+06	3.88E+05	9.68E+05	3.32E+06	5.26E+06	8.77E+06	1.23E+07	1.45E+07
4	4,5	8.47E+06	3.73E+05	8.76E+05	3.04E+06	4.84E+06	8.07E+06	1.13E+07	1.34E+07

As you can see in Table C2 with regard to first policy and viewpoint of customers, plan 4 (hubs 4 and 5) is the best plan, because in first policy the mitigating customers' cost is very high priority. In the viewpoint of constructors though, the best plant is plan 1 (hubs 1 and 7) and the worst plan is plan four. Please note that if constructors are persuaded to select plan 4 instead plan 1, they lose 28.71 percent of their benefits (5383870) but instead customers' cost reduced by 6.5 present (5527060). As you see in the classical hub problem, plan 4 will be selected but in real-world, constructors highly tend to choose plan 1.

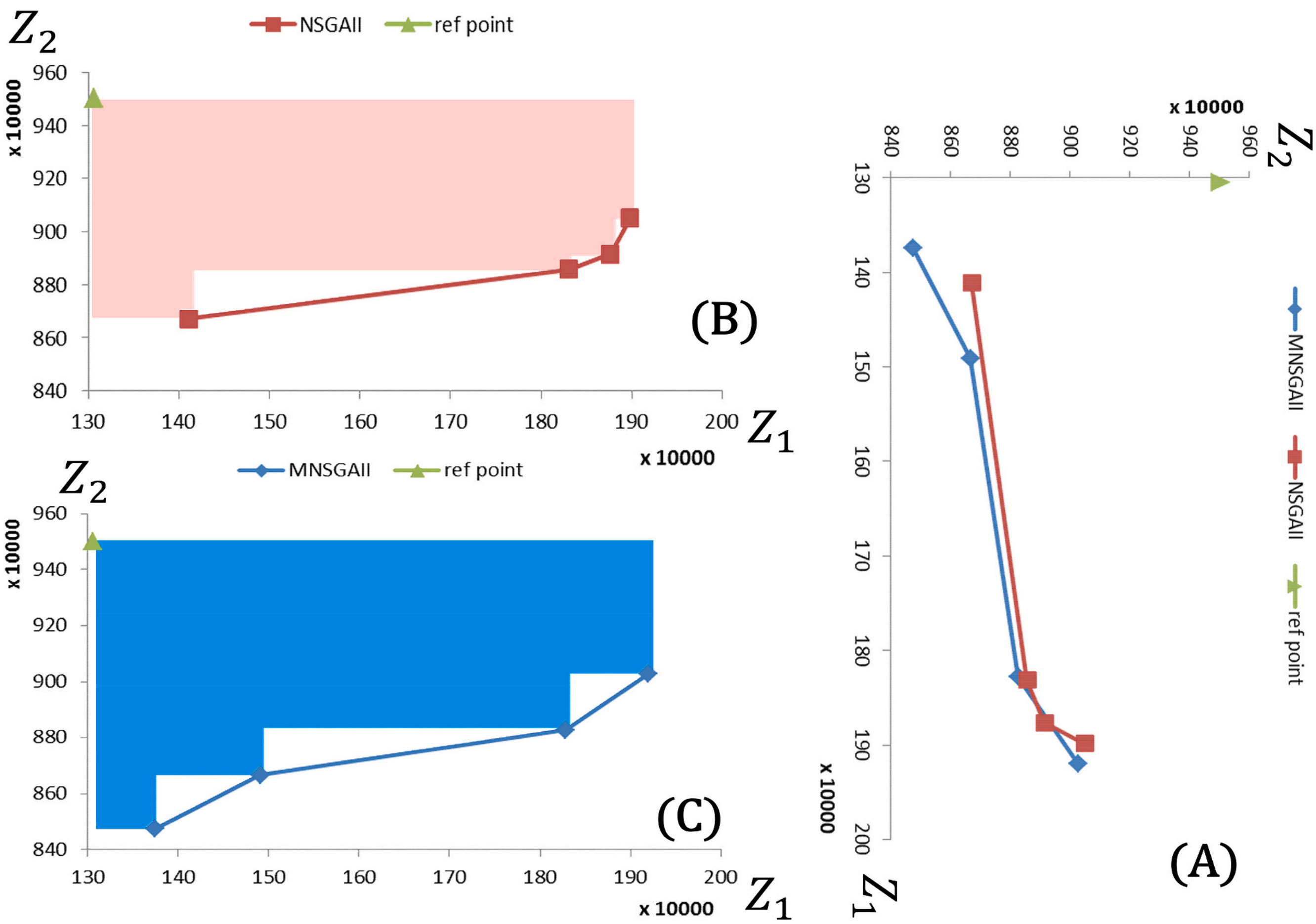


Fig. C.4. Approximated Pareto Front of illustrative example.

Appendix D. The modified orthogonal array L_{18}

# trail	Factors						
	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	2	2	2	2	2	2
3	1	3	3	3	3	3	3
4	2	1	1	2	2	3	3
5	2	2	2	3	3	1	1
6	2	3	3	1	1	2	2
7	3	1	2	1	3	2	3
8	3	2	3	2	1	3	1
9	3	3	1	3	2	1	2
10	1	1	3	3	2	2	1
11	1	2	1	1	3	3	2

(continued on next page)

(*continued*)

# trail	Factors						
	1	2	3	4	5	6	7
12	1	3	2	2	1	1	3
13	2	1	2	3	1	3	2
14	2	2	3	1	2	1	3
15	2	3	1	2	3	2	1
16	3	1	3	2	3	1	2
17	3	2	1	3	1	2	3
18	3	3	2	1	2	3	1

References

Alkaabneh, F., Diabat, A., & Elhedhli, S. (2019). A Lagrangian heuristic and GRASP for the hub-and-spoke network system with economies-of-scale and congestion. *Transportation Research Part C: Emerging Technologies, 102*, 249–273. https://doi.org/10.1016/j.trc.2018.12.011.

Alumur, S., & Kara, B. Y. (2008). Network hub location problems: The state of the art. *European Journal of Operation Research, 190*, 1–21. https://doi.org/10.1016/j.ejor.2007.06.008.

Atashpaz-Gargari, E., & Lucas, C. (2007). Imperialist competitive algorithm: An algorithm for optimization inspired by imperialistic competition. *IEEE Congress on*. https://doi.org/10.1109/CEC.2007.4425083. Evolutionary Computation, 2007. CEC 2007.

Barnes, G., & Langworthy, P. (2003). *The Per-Mile Costs OF Operating Automobiles and Trucks*. St. Paul, MN 55155: Minnesota Department of Transportation Office of Research Services, 395 John Ireland Boulevard.

Bashiri, M., Mirzaei, M., & Randall, M. (2013). Modeling fuzzy capacitated p-hub center problem and a genetic algorithm solution. *Applied Mathematical Modelling*, 3513–3525. https://doi.org/10.1016/j.apm.2012.07.018.

Bernal, E., Castillo, O., Soria, J., & Valdez, F. (2017). Imperialist competitive algorithm with dynamic parameter adaptation using fuzzy logic applied to the optimization of mathematical functions. *Algorithms, 10*(18).

Bernal, E., Castillo, O., Soria, J., & Valdez, F. (2020). Parameter adaptation in the imperialist competitive algorithm using generalized type-2 fuzzy logic. *Intuitionistic and Type-2 Fuzzy Logic Enhancements in Neural and Optimization Algorithms: Theory and Applications*, 3–10.

Campbell, J. (1996). Hub location and the p-hub median problem. *Operations Research, 44*(6), 1–13. https://doi.org/10.1287/opre.44.6.923.

Cedex. (2010). *Guide ON Economic Evaluation OF Transport Projects*. Centro de Estudios y Experimentación de Obras Públicas (CEDEX).

Commission, E. (1997). *Vehicle Taxation IN the European Union 1997*. Brussels: EUROPEAN COMMISSION.

Correia, I., Nickel, S., & Saldanha-da-Gama, F. (2018). A stochastic multi-period capacitated multiple allocation hub location problem: Formulation and inequalities. *Omega, 74*, 122–134. https://doi.org/10.1016/j.omega.2017.01.011.

DaCoTA. (2012). *Vehicle safety. Deliverable 4*. 8u of the EC FP7 project DaCoTA.

Damgacioglu, H., Dinler, D., Ozdemirel, N. E., & Iyigun, C. (2014). A genetic algorithm for the uncapacitated single allocation planar hub location problem. *Computers & Operations Research*. https://doi.org/10.1016/j.cor.2014.09.003.

Damgacioglu, H., Dinler, D., Ozdemirel, N. E., & lygun, C. (2014). A genetic algorithm for the uncapacitated single allocation planar hub location problem. *Computers & Operations Research*. https://doi.org/10.1016/j.cor.2014.09.003.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation, 6*(2), 182–197. https://doi.org/10.1109/4235.996017.

Ebrahimi Zade, A., Sadegheih, A., & Lotfi, M.m. (2014). A modified NSGA-II solution for a new multi-objective hub maximal covering problem under uncertain shipments. *Journal of Industrial Engineering International, 10*(4), 185–197. https://doi.org/10.1007/s40092-014-0076-4.

Ernst, A., & Krishnamoorthy, M. (1996). Efficient algorithms for the uncapacitated single allocation p-hub median problem. *Location Science, 4*(1), 39–54.

European Economy. (2013). Tax reforms in EU Member States, Tax policy challenges for economic growth and. *European Economy, 5*.

Gelareh, S., & Nickel, S. (2011). Hub location problems in transportation networks. *Transportation Research Part E: Logistics and Transportation Review, 47*(6), 1092–1111. https://doi.org/10.1016/j.tre.2011.04.009.

Ghezavati, V., & Hosseinifar, P. (2018). Application of efficient metaheuristics to solve a new bi-objective optimization model for hub facility location problem considering value at risk criterion. *Soft Computing, 22*(1), 195–212. https://doi.org/10.1007/s00500-016-2326-4.

Ilić, A., Urošević, D., Brimberg, J., & Mladenović, N. (2010). A general variable neighborhood search for solving the uncapacitated single allocation p-hub median problem. *European Journal of Operational Research, 206*(2), 289–300.

Indian roads congress. (2009). *Manual on economic evaluation of highway projects in india (Second Revsion ed.)*. New Delhi: Indian roads congress.

Kimms, A. (2005). *Economies of scale in hub & spoke network design models: We have it all wrong*. Technical Report (Vol. 45). Lessingstrasse: Freiberg, Germany: Technische Universitat Bergakademie Freiberg.

Klincewicz, J. (2002). Enumeration and search procedures for a hub location problem with economies of scale. *Annals of Operations Research, 110*, 107–122.

Ko, S., Lautala, P., Fan, J., & Shonnard, D. (2019). Economic, social, and environmental cost optimization of biomass transportation: A regional model for transportation analysis in plant location process. *Biofuels, Bioproducts Bioreference, 13*, 582–598. https://doi.org/10.1002/bbb.1967.

Lüer-Villagra, A., Eiselt, H., & Marianov, V. (2019). A single allocation p-hub median problem with general piecewise-linear costs in arcs. *Computers & Industrial Engineering, 128*, 477–491. https://doi.org/10.1016/j.cie.2018.12.058.

LuerVillagra, A., & Marianov, V. (2013). A competitive hub location and pricing problem. *European Journal of Operational Research, 231*(3), 734–744. https://doi.org/10.1016/j.ejor.2013.06.006.

Martí, R., Corberán, Á., & Peiró, J. (2015). Scatter search for an uncapacitated p-hub median problem. *Computers & Operations Research, 58*, 53–66. https://doi.org/10.1016/j.cor.2014.12.009.

O`Kelly, Morton, E., & Harvey, J. (1994). The hub network design problem: A review and synthesis. *Journal of Transport Geography, 2*(1), 31–40.

Panneerselvam, R. (2013). *Engineering Economics*. Delhi: PHI Learning Pvt. Ltd.

Pham, D., & Karaboga, D. (2000). *Intelligent optimization Techniques:genetic algorithms, Tabu search, simulated Annealing and Neural networks*. London: Springer-Verlag.

Sasaki, M., Campbell, J., Krishnamoorthy, M., & Ernst, A. (2014). A Stackelberg hub arc location model for a competitive environment. *Computers & Operations Research, 47*, 27–41. https://doi.org/10.1016/j.cor.2014.01.009.

Taguchi, G. (1986). *Introduction to quality Engineering: Designing quality into products and Processes*.

Tavakkoli-Moghaddam, R., Gholipour-Kanani, Y., & Shahramifar, M. (2013). A multi-objective imperialist competitive algorithm for a capacitated single-allocation hub location problem. *International Journal of Engineering, 26*(6), 605–620.

Teck Sim, T. K. (2007). *The hub covering flow problem and the stochastic p-hub center problem*. University of Iowa.

Wagner, B. (2004). *Model formulation for hub covering problems. Working paper. Hochschulstrasse 1, 64289 Darmstadt, Germany*. Instittute of Operation Research, Darmstadt University of Technology.

Wu, J., & Azarm, S. (2000). Metrics for quality assessment of a multiobjective design optimization solution set. *Journal of Mechanical Design, 123*(1), 18–25. https://doi.org/10.1115/1.1329875.

Yaman, H. (2011). Allocation strategies in hub networks. *European Journal of Operational Research*, 442–451.

Zanjirani Farahani, R., Hekmatfar, M., Boloori Arabani, A., & Nikbakhsh, E. (2013). Hub location problems: A review of models, classification, solution techniques, and applications. *Computers & Industrial Engineering, 64*, 1096–1109. https://doi.org/10.1016/j.cie.2013.01.012.

Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., & Da Fonseca, V. G. (2003). Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on Evolutionary Computation, 7*(2), 117–132. https://doi.org/10.1109/TEVC.2003.810758.