Noname manuscript No.
(will be inserted by the editor)

A Genetic Algorithm for solving Bus Terminal Location problem using Data Envelopment Analysis with Multi Objective Programming

Atefeh Taghavi, Reza Ghanbari, Khatere Ghorbani-Moghadam, Alireza Davoodi, Ali Emrouznejad

the date of receipt and acceptance should be inserted later

Abstract Due to urbon expansion and population increasing bus network design is an important problem in public transportation. Functional aspect of bus networks such as fuel consumption and the depreciation of buses and also spatial aspects of bus networks such as station and terminal locations or access rate to the buses are not proper conditions in most cities. Therefore, having an efficient method to evaluate the performance of bus lines considering both functional and spatial aspects is essential. In this paper, we propose a new model for bus terminal location problem (BTLP) using data envelopment analysis (DEA) with multi-objective programming approach. In this model, we want to find an efficient allocation patterns for assigning stations terminals and also we find the optimal locations for deploying terminals. So we use genetic algorithm for solving our model. By simultaneous combing the data envelopment analysis and bus terminal location problem, two types of

A. Taghavi
Faculty of Mathematical Sciences, Department of Applied Mathematics, Ferdowsi University of Mashhad, Mashhad, Iran.
E-mail: at.taghavi@yahoo.com

Kh. Ghorbani-Moghadam
Member of Optimization Laboratory in Faculty of Mathematical Sciences, Department of Applied Mathematics, Ferdowsi University of Mashhad, Mashhad, Iran.
E-mail: khatere.moghadam67@gmail.com

R. Ghanbari
Faculty of Mathematical Sciences, Department of Applied Mathematics, Ferdowsi University of Mashhad, Mashhad, Iran.
E-mail: rghanbari@um.ac.ir

A. Davoodi
Department of Mathematics, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran.
E-mail: alirzd@yahoo.com.

A. Emrouznejad
Aston Business School, Aston University, Birmingham, Uk.
E-mail: a.emrouznejad@aston.ac.uk.

efficiencies are optimized: spatial efficiency as measured by finding allocation patterns with the most serving amount, and the terminals efficiency in serving demands as measured by the data envelopment analysis efficiency score for selected allocation patterns. This approach is useful when terminals efficiency are one of the important criteria in choosing the optimal terminals location for decision makers.

Keywords Bus terminal location problem · Data envelopment analysis · Efficiency · Genetic Algorithm · Multi objective programming.

1 Introduction

From long time ago urban transport plays an essential role in city development. Therefore, one important factor of urbanization and city development is proper access to public transportation. Among all urban transportations, bus transit is more predominant and it is distinct form others due to low investment and providing a lot of services. Since bus lines have the most significant role in our public transportation, they should have the necessary efficiency and provide the best services for citizens. These services should be safe, cheap, fast and desirable. Consequently, optimal and efficient bus network design to improve public transportation conditions is very important. One main step in this way is to distinguish optimal numbers and locations of bus terminals so that we also have efficient allocation patterns to allocate demand bus stations to terminals. Here, we pay specific attention for BTLP and DEA. By simultaneous combing of these two problems, the effect of terminal locations on efficiency of services to the stations is considered. BTLP considered as a special case of facility location problem, so it is NP-hard [22]. Its exact solution for big cities is usually complicated and time-consuming. Therefore, we decide to design a meta-heuristic algorithm based on genetic to solve the combined model. Finally, we study the performance of our proposed algorithm by evaluating the numerical results.

The remainder of our work is organized as follows: In Section 2, describes a mathematical model of BTLP and we provide a brief introduction to DEA models. In Section 3, we develop and present the combined model of BTLP with the DEA problem. In Section 4, we explain the proposed genetic algorithm for solving combined model of BTLP and DEA problem. In Section 5, we apply these models to a small hypothetical data set and present the numerical results. Finally, In Section 6, conclusions and future extensions are discussed.

2 Preliminaries

2.1 An introduction to the bus terminal location problem

One of a significant problem in design of bus system is to determine the suitable location to construct bus terminals [2,3,4,12,13,16,18,27,29,34]. Consider a set of public transportation stations like bus station while the number of passengers of each station and the distance of each station to the candidate terminal are given. The purpose of BTLP is to determine the locations of the required terminals and allocate demand stations to them, so that public service is maximized by the construction of the terminals. In order to have a proper transportation system, passengers of each station are allocated to the nearest constructed terminal. According to conditions of the problem and researches in [27], three neighborhoods are defined. For each terminal, and for each neighborhood is considered a terminal service desirability to the stations. This model is a kind of correction to the model provided by Ghanbari and Mahdavi-Amiri [18]. Ghanbari and Mahdavi-Amiri [18] defined a neighborhood for each terminal so that if they are established, they will service all of their available stations in the neighborhood, and the service desirability is a decreasing function based on the distance of the station from the desired terminal. In fact, in the model provided by Ghanbari and Mahdavi-Amiri [18], the service desirability is affected by distance, although the difference distance between the stations from the terminal is partial. According to the studies done in [27], it was concluded that if the difference distance between the stations from the terminal is partial, then the amount of satisfaction of the stations from the terminal is the same. This means that if the station's distance from the terminal is less than a constant amount, then the service desirability of the terminal to the stations are equal to each other. So, for each terminal, we define three neighborhoods so that the service desirability of the terminal to the available stations in the neighborhood is assumed to be a constant number. In other words, fixing this service desirability, means that if the distance between a station and a number of terminals is less than the constant number, there will be no difference as to the station receives service from which terminal. Regarding this, in contrast to the model provided by Ghanbari and Mahdavi-Amiri [18], three types of neighborhoods are defined for each terminal.
First, we define data and neighborhoods before expressing the model of BTLP.

2.1.1 Data definition [18]

We consider that m and n denote the number of candidate centers for bus terminals and the number of demand bus stations, respectively. We define $K = \{1, ..., m\}$ and $L = \{1, ..., n\}$. We also consider that $C = [c_{kl}]$ is a distance matrix between node $k \in K$ and $l \in L$ and $F = \{f_1, f_2, ..., f_n\}$ is a set of passengers (potentials) corresponding to each node l (more explicitly, d_l equals to the maximum number of passengers corresponding to each $l \in L$).

The objective is to select p terminals from K so that the public service function is maximized.

2.1.2 Neighborhood definition

Definition 1 For each $k \in K$, we have define three neighborhoods [27]:

$$L^*_{1k} = \{l \in L; c_{kl} < r_1\},$$
$$L^*_{2k} = \{l \in L; r_1 \leq c_{kl} < r_2\},$$
$$L^*_{3k} = \{l \in L; c_{kl} \geq r_2\}.$$

Where, $r_1 < r_2 \in \mathbb{R}^+$ are constant and they are called neighborhood radius.

Definition 2 For each l, we have:

$$K^*_{1l} = \{k \in K; l \in L^*_{1k}\},$$
$$K^*_{2l} = \{k \in K; l \in L^*_{2k}\},$$
$$K^*_{3l} = \{k \in K; l \in L^*_{3k}\}.$$

For stations in the neighborhood of the first type (L^*_{1k}), service desirability is considered constant number C_1. Neighborhood type two (L^*_{2k}) is also like neighborhood type one. But there is a difference and that is service desirability terminal to the stations are less than C_1 and is constant number C_2 since its stations are farther away as compared with stations of neighborhood type one. Neighborhood type three (L^*_{3k}), is considered that if the stations distance to the terminal is greater than constant (r_2), the service desirability is a decreasing function of station distance from the terminal same as defined in [18]. In other words, if the distance is greater than constant (r_2) servicing amount is affected by the distance.

2.1.3 Variable definition

x_{kl} and y_k are binary variables. We define them as follows:

$$x_{kl} = \begin{cases} 1, \text{If node } l \text{ receives services from terminal } k, \\ 0, \text{otherwise.} \end{cases}$$

$$y_k = \begin{cases} 1, \text{If candidate center } k \text{ is selected as a terminal,} \\ 0, \text{otherwise.} \end{cases}$$

We define BTLP as follows:

$$max \quad \sum_{k\in K}\left(\sum_{l\in L^*_{1k}} C_1 f_l x_{kl} + \sum_{l\in L^*_{2k}} C_2 f_l x_{kl} + \sum_{l\in L^*_{3k}} \frac{1}{c_{kl}} f_l x_{kl}\right) \tag{1·2}$$

$$s.t.$$

$$\sum_{l\in L^*_{1k}} x_{kl} + \sum_{l\in L^*_{2k}} x_{kl} + \sum_{l\in L^*_{3k}} x_{kl} \le |L| y_k, \qquad \forall k \in K \tag{2·2}$$

$$\sum_{k\in K^*_{1l}} x_{kl} + \sum_{k\in K^*_{2l}} x_{kl} + \sum_{k\in K^*_{3l}} x_{kl} = 1, \qquad \forall l \in L \tag{3·2}$$

$$\sum_{k=1}^{m} y_k = p, \tag{4·2}$$

$$x_{kl}, y_k \in \{0,1\}. \qquad \forall k, l \tag{5·2}$$

The objective function (1.2) determines the servicing amount of the candidate nodes to all nodes. Constraint (2.2) shows only if serving center k is selected a terminal, it can service to the nodes around. Constraint (3.2) emphasizes that each node l only serves from a terminal. Constraint (4.2) ensures that the number of terminals required is not more than p.

2.2 Data envelopment analysis

Recent years, in most countries there are different application of data envelopment analysis to evaluate institution performances and other common activities in different fields. In fact, DEA is a linear programming technique useful to assess relative efficiency among similar entities that in DEA are called Decision Making Units (DMUs). It was introduced by Charnes, Cooper and Rhodes [7] based on Farrell [15] pioneering work. They generalized the single-output to single-input ratio definition of efficiency to multiple inputs and outputs. In their original DEA model, Charnes, Cooper and Rhodes (CCR model) proposed that the efficiency of a DMU can be obtained as the maximum of a ratio of weighted outputs to weighted inputs, subject to the condition that the same ratio for all DMUs must be less than or equal to one. To evaluate a DMU, the best possible set of weights for that particular DMU is selected and the efficiency score of all DMUs in the set are computed using those weights. If the DMU being evaluated obtains the highest score, it is classified as efficient, otherwise as inefficient. The efficient DMUs generate an efficient frontier that envelops all DMUs.

In addition, in DEA methods unlike some of numerical methods, it is not necessary to have predetermined weights and allocate them to the inputs and outputs. Additionally, these methods do not need predetermined forms functions (such as statistical regression methods) or exact form of production function (such as some parametric methods). DEA method is based on linear algebra and its ability is because of using the solving methods of linear programming problem. Therefore, DEA determines source and the amount of inefficiency for each input and output [10]. See other works on DEA method in [8, 26, 31].

2.2.1 The modeling of DEA problem

Let v_i, $i = 1, \ldots, I$ and u_j, $j = 1, \ldots, J$ being inputs and outputs respectively. We should solve all k optimization problems wich to be able to evaluate a DMU each time. We calculate weights by using v_i and u_j. We solve the following fractional problem for r-th DMU (DEA 1) [7]:

$$max \quad \frac{\sum_{j=1}^{J} u_j O_{jr}}{\sum_{i=1}^{I} v_i I_{ir}} \tag{6 $\cdot$ 2}$$

$$s.t.$$

$$\frac{\sum_{j=1}^{J} u_j O_{jk}}{\sum_{i=1}^{I} v_i I_{ik}} \leq 1, \qquad \forall k \tag{7 $\cdot$ 2}$$

$$u_j, v_i \geq 0. \qquad \forall j, i \tag{8 $\cdot$ 2}$$

Where, O_{jk} and I_{ik} are the amount of j-th output and i-th input for k-th DMU. These constraints guarantee that the ratio of sum of the weighted outputs to sum of the weighted inputs is not greater than one for none of DMUs. The achieved v_is and u_js maximize the efficiency for r-th DMU because of the objective function form. Similar programs are successively solved for other DMUs to calculate the optimal efficiency and sum of the weights for each DMU.

The denominator of objective function (6.2) is assumed one to solve the fractional model as a linear programming ($\sum_{i=1}^{I} v_i I_{ir} = 1$) and it is considered as a constraint. So we maximize the numerator of fraction (6.2) and by multiplying both sides of constraint (7.2) by the denominator, i.e. sum of the weighted inputs, we have the following linear constraints:

$$\sum_{j=1}^{J} u_j O_{jk} \leq \sum_{i=1}^{I} v_i I_{ik}, \qquad \forall k.$$

In addition, to prevent the creation of weak efficiency for DMUs in linear programming of DEA, it is necessary that every weight is greater than $\varepsilon > 0$ (a small infinite amount). The weak efficiency for a special DMU happens when the achieved optimal amount of efficiency (w^*) from DEA model becomes one and at least one of the weights of input and output (v^*, u^*) in optimal solution equals to zero.

Therefore, the fractional model of DEA for the r-th DMU is converted to the following linear programming (DEA 2) [7]:

$$max \quad \sum_{j=1}^{J} u_j O_{jr} \tag{9 · 2}$$

$$s.t.$$

$$\sum_{i=1}^{I} v_i I_{ir} = 1, \tag{10 · 2}$$

$$\sum_{j=1}^{J} u_j O_{jk} - \sum_{i=1}^{I} v_i I_{ik} \leq 0, \qquad \forall k \tag{11 · 2}$$

$$u_j, v_i \geq \varepsilon. \qquad \forall j, i \tag{12 · 2}$$

Note that for efficiency evaluation of each DMU, (DEA 2) optimization problem should be solved k times. Each problem has $I + J$ decision variables and $K + 1$ constraints.

2.2.2 The Simultaneous of the DEA model

In DEA, the DMUs should be equivalence. It means that they should have similar inputs and outputs. Sometimes, because of lots of DMUs, many linear programming models should be written and the process of problem solving is time-consuming. To solve these problems, Klimberg [20] presented a model that it has a variable named d_r, the inefficiency level of the r-th DMU, and the input d_r in objective function ($d_r = 1 - w_r$). Therefore, for constraint (11.2) in linear programing model of DEA, we have:

$$\sum_{j=1}^{J} u_j O_{jk} - \sum_{i=1}^{I} v_i I_{ik} + d_r = 0.$$

Using constraint (10.2) for the above relation, for DMUr we have:

$$\sum_{j=1}^{J} u_j O_{jr} + d_r = 1.$$

Then, we have the following modified data envelopment ananlysis model (MDEA2) [20]:

$$
\begin{aligned}
& max \quad w_r = 1 - d_r && && (13 \cdot 2)\\
& s.t. && && \\
& \sum_{i=1}^{I} v_i I_{ir} = 1, && && (14 \cdot 2)\\
& \sum_{j=1}^{J} u_j O_{jr} + d_r = 1, && && (15 \cdot 2)\\
& \sum_{j=1}^{J} u_j O_{jk} - \sum_{i=1}^{I} v_i I_{ik} \leq 0, && \forall k \neq r && (16 \cdot 2)\\
& u_j, v_i \geq \varepsilon. && \forall j, i && (17 \cdot 2)
\end{aligned}
$$

We can expand MDEA2 model for all units simultaneously. This model is named simultaneous data envelopment analysis model (SDEA) and it is as follows [20]:

$$
\begin{aligned}
& max \quad \sum_r w_r = \sum_r (1 - d_r) && && (18 \cdot 2)\\
& s.t. && && \\
& \sum_{i=1}^{I} v_{ri} I_{ir} = 1, && \forall r && (19 \cdot 2)\\
& \sum_{j=1}^{J} u_{rj} O_{jr} + d_r = 1, && \forall r && (20 \cdot 2)\\
& \sum_{j=1}^{J} u_{rj} O_{jk} - \sum_{i=1}^{I} v_{ri} I_{ik} \leq 0, && \forall k, r; k \neq r && (21 \cdot 2)\\
& d_r, u_{rj}, v_{ri} \geq \varepsilon. && \forall j, i, r && (22 \cdot 2)
\end{aligned}
$$

Where, d_r, u_{rj} and v_{ri} are decision variables to show inefficiency level of the r-th DMU, the j-th output weight and the i-th input weight.

3 combined model of DEA and BTLP

As mentioned before, one of the major socio-economic activities in modern urban societies is transportation. The lack of an efficient transportation system is one of the obstacles to the growth and development of any country. Given the fact that among the different types of urban transport, the bus network has

the largest role and share, so it must be efficient and able to provide the most and best service to citizens. As a result, evaluating the efficiency of the bus network is of special importance. For this reason, many studies have been carried out on the efficiency of public transportation. In 2012, Holmgren [19] was studied with the aim of evaluating the efficiency of public transport operations undertaken in swedish counties by the public transport authorities, taking in to account the substantial differences in operating conditions between counties with using annual data from 1986 to 2009 for 26 Swedish counties. It is concluded that the efficiency of the public transport providers in all counties fell during the observed time period from 2000 to 2009. Finally, the research focuses on increased emphasis on route density as well as higher environmental and safety requirements [19]. In 1997, Viton [33] evaluated the efficiency of the U.S. bus system by using the DEA method. This research studies the efficiency of U.S. multi-mode bus transit systems by asking whether they could expand their service (outputs) without requiring additional resources (inputs); or whether they could reduce input utilization without having to reduce service. At the end the findings indicate that 80 of the bus systems studied are technically efficient, and that the extent of inefficiency in the industry is slight [33]. In 2009, Lao and Liu [23] used DEA and geographic information systems (GIS) to evaluate the performance of bus lines within a public transit system, considering both the operations and operational environment. This study suggests ways to improve the performance of bus lines [23].
As you can see in the mentioned research, in most of them, the functional aspect of bus network is only considered and spatial components are considered less. In appropriate allocation of terminals and stations make extra expenses and decrease the bus users. The difference between this paper and other researches is the efficiency evaluation of terminals and stations considering functional and spatial aspects so that besides determining the optimal locations patterns for assigning stations to terminals are also determined.

In this section, we state the combined model of bus terminal location and DEA problem with one input and two outputs. To understand more, we first explain input and outputs in DEA model. We consider the distance traveled by terminal buses that service the covered stations for input. Note that when this distance is lesser, the amount of fuel consumption will be lesser. In addition, the number of routs with special lines and the number of BRT buses are considered as outputs indicators. Parameter and decision variables are given before expressing the combined model.

Problem parameters:

Decision variables:

c_{kl} The distance between terminal k and station l.

f_l Sum of passengers in station l.

I_{kl} The amount of distance traveled by buses of terminal k for assigning to station l.

O_{kl1} The number of routs with special lines for assigning terminal k to station l.

O_{kl2} The number of BRT buses for assigning terminal k to station l.

x_{kl} Binary variable for Service or unavaible service of the constructed terminal in center k to station l.

y_k Binary variable for constructing or not constructing a terminal in center k.

d_{kl} a none negative variable for inefficiency level for assigning terminal k to station l.

v_{kl} a none negative variable for input weight for assigning terminal k to station l.

u_{kl1} a none negative variable for weight of output 1 for assigning terminal k to station l.

u_{kl2} a none negative variable for weight of output 2 for assigning terminal k to station l.

Now, the combined model is as follows:

$$max \quad \sum_{k=1}^{K}\sum_{l=1}^{L}(1-d_{kl}) \tag{1 $\cdot$ 3}$$

$$max \quad \sum_{k\in K}\left(\sum_{l\in L^*_{1k}} C_1 f_l x_{kl} + \sum_{l\in L^*_{2k}} C_2 f_l x_{kl} + \sum_{l\in L^*_{3k}} \frac{1}{c_{kl}} f_l x_{kl}\right) \tag{2 $\cdot$ 3}$$

$s.t.$

$$\sum_{l\in L^*_{1k}} x_{kl} + \sum_{l\in L^*_{2k}} x_{kl} + \sum_{l\in L^*_{3k}} x_{kl} \le |L| y_k, \qquad \forall k\in K, \tag{3 $\cdot$ 3}$$

$$\sum_{k\in K^*_{1l}} x_{kl} + \sum_{k\in K^*_{2l}} x_{kl} + \sum_{k\in K^*_{3l}} x_{kl} = 1, \qquad \forall l\in L, \tag{4 $\cdot$ 3}$$

$$\sum_{k=1}^{m} y_k = p, \tag{5 $\cdot$ 3}$$

$$v_{kl} I_{kl} = x_{kl}, \qquad \forall k, l, \tag{6 · 3}$$

$$u_{kl1} O_{kl1} + u_{kl2} O_{kl2} + d_{kl} = 1, \qquad \forall k, l, \tag{7 · 3}$$

$$u_{kl1} O_{rs1} + u_{kl2} O_{rs2} - v_{kl} I_{rs} \leq 0, \qquad \forall k, l, r, s; k \neq r, l \neq s, \tag{8 · 3}$$

$$u_{kl1} \geq \varepsilon x_{kl}, \qquad \forall k, l, \tag{9 · 3}$$

$$u_{kl2} \geq \varepsilon x_{kl}, \qquad \forall k, l, \tag{10 · 3}$$

$$v_{kl} \geq \varepsilon x_{kl}, \qquad \forall k, l, \tag{11 · 3}$$

$$u_{kl1} O_{kl1} \leq x_{kl}, \qquad \forall k, l, \tag{12 · 3}$$

$$u_{kl2} O_{kl2} \leq x_{kl}, \qquad \forall k, l, \tag{13 · 3}$$

$$x_{kl}, y_k \in \{0, 1, \} \qquad \forall k, l, \tag{14 · 3}$$

$$d_{kl}, u_{kl1}, u_{kl2}, v_{kl} \geq 0. \qquad \forall k, l, \tag{15 · 3}$$

If a station is serviced by terminal, according to inputs and outputs amounts we have an efficiency score. It means that for each pair terminal k and station l that $x_{kl} = 1$, we have an efficiency score. The objective function (1.3), maximizes the sum of the efficiencies for all terminal K and station L combinations that may exist in the optimal solution. Objective function (2.3) and constraints (3.3) to (5.3) are the objective function and constraints of the BTLP model. In addition, when terminal k services station l ($x_{kl} = 1$), constraints (6.3) to (11.3) are the same as constraints of the SDEA. On other hand, if terminal k does not serve station l ($x_{kl} = 0$), the constraints (9.3) to (11.3) require the input and output weights to be non-negative, and the constraint in (6.3), (12.3) and (13.3) force them to be equal to 0.
This bio objective combined model has $5KL + K$ decision variables and $1 + K + L + 7KL + KL(K-1)(L-1)$ constraints (without considering sign and one and zero constraints) and it works like this that in each execution of search algorithm, that used to solve this model, of solution region, an allocation-location pattern is selected so that the constraints are satisfied. It means that a feasible solution is selected. Then the cost of this pattern is computed. Afterwards, for each not selected allocations, d_{kl} equals one. Therefore, the first objective function is zero for not selected allocations. In addition for selected allocations, first objective function is as follows:

$$\sum_{k=1}^{K} \sum_{l=1}^{L} (1 - d_{kl}) = \sum_{k=1}^{K} \sum_{l=1}^{L} (x_{kl} - d_{kl}) = \sum_{k=1}^{K} \sum_{l=1}^{L} \sum_{j=1}^{J} u_{klj} O_{klj}.$$

It is sum of the weighted outputs. In fact, for the selected allocations, a DEA model is solved. In the other words, the relative efficiency of the selected allocations are computed to each other. The best selected allocations make the efficient frontier and the efficiency of other selected allocations are computed in proportion to this efficient frontier. This process continues until a pattern is chosen with minimum cost and maximum efficiency. We can name the performance of above model simultaneous combine of location and DEA model. It means that at each iteration of search algorithm, a DEA model is simultaneously solved and the efficiency of the selected allocations are computed.

4 The proposed genetic algorithm to solve the combined of DEA and BTLP

In this section we want to propose a genetic algorithm for solving combined model of BTLP and DEA problem. So the necessary preliminaries are presented before the proposed genetic algorithm.

Definition of solution region

Suppose $K = \{1, ..., m\}$ is the index set of the candidate terminals. For each feasible solution y, K is partitioned in to two distinct sets; the set of close terminals (K_c) and the set of open terminals (K_o). It means:

$$K_o = \{k|y_k = 1\} \ , \ K_c = \{k|y_k = 0\}.$$

If S ($S = K_o$) is shown the solution set, i.e. the points that terminals should be built on; S is feasible if $S \subset K$ and $|S| = P$. We show the set of all feasible solutions of BTLP by Y.

Initialization

Population is initialized by random numbers zero and one. Each chromosome that shows one y, is made by random numbers zero and one so that $y \in Y$.

Selection

The selected operator uses tournament selection in this algorithm, i.e. members are selected according to the ranking of fitness function. The amounts of fitness functions are arranged in descending order since DEA and BTLP models are maximization ones and as a result members with maximum fitness are selected.

Crossover

Crossover operator is done for two chromosomes selected randomly and two new children entered in to population this operator acts in a way that for each similar components in two chromosomes, children are initialized by those similar components. In addition for components that two chromosomes are different, children inherit randomly some components from a chromosome and other components from other chromosome so that the number of component one are exactly P.

Mutation

Mutation operator acts in a way that one y is selected randomly from population since the number of open terminals are P. Then two of its components are randomly identified. If one component is one and the other is zero, the components are replaced. So, the number of ones do not change.

Stop condition

The algorithm is terminated after arbitrary number R generation.

Fitness function

To calculate fitness function of a solution, first the normalized weights w_1 and w_2 $(w_1+w_2=1)$ are considered. w_1 and w_2 are the weights of BTLP objective function (f_1 in (1.4)) and DEA objective function (f_2 in (2.4)) , respectively. So that their amounts are arbitrary values in $[0,1]$. Therefore, fitness function is defined as follows:

$$f_1 = \sum_{k\in K}\left(\sum_{l\in L^*_{1k}} C_1 f_l x_{kl} + \sum_{l\in L^*_{2k}} C_2 f_l x_{kl} + \sum_{l\in L^*_{3k}} \frac{1}{c_{kl}} f_l x_{kl}\right), \qquad (1\cdot 4)$$

$$f_2 = \sum_{k=1}^{K}\sum_{l=1}^{L}(1-d_{kl}), \qquad (2\cdot 4)$$

$$fitness = w_1 f_1 + w_2 f_2. \qquad (3\cdot 4)$$

Generally, in this algorithm, members of population (y) are initialized randomly. Then for each member, matrix x that shows available or unavailable service of open terminals to the stations, is calculated according to the shortest stations distances from open terminals. Therefore, the amount of f_1 is distinguished. Afterwards, for each y and x, a linear optimization model of DEA is solved and input and outputs weights and the amount of f_2 for each y are determined. Finally, amount of fitness is computed with arbitrary w_1 and w_2. Next, they are ranked by section operator of fitness functions. In addition, crossover and mutation operators are applied to population and new children are generated. Children are added to the current population and members with worst fitness are deleted from population. Now, the new improved population is transferred to the next generation. Repeating the algorithm makes the fitness of population members and consequently the solutions be improved.
Before explanation of steps of this process in algorithm 1, we introduce the symbols used in the algorithm:

Y	Set of members in population.
y	A member in population Y.
M	Size of population.
R	Member of generation going to be produced.
Y_r	r-th generation of population.

Algorithm 1 The proposed genetic algorithm

1 Initialization: Get the values M and R.

2 Put $r \leftarrow 1$.

3 Generate the initial population Y_r randomly.

4 Repeat until $r \leq R$:

  4-1 Repeat for each member of the population ($y \in Y_r$):

    4-1-1 Calculate the values of matrix x.

    4-1-2 Calculate the f_1 by using equation (1.4) .

    4-1-3 Solve the DEA linear model and calculate the f_2 by using equation (2.4).

    4-1-4 Calculate the fitness by using equation (3.4).

  4-2 Use the selection operator on the population.

  4-3 Use the Crossover operator on the selected pairs.

  4-4 Use the Mutation operator.

  4-5 Replace part of the current population with the best children.

  4-6 $r \leftarrow r + 1$.

  4-7 Show the best member of the current population in the output.

5 Numerical results

In this section, the numerical results of the proposed genetic algorithm on a set of random data with 8 members in population are given. To test our proposed algorithm, we create examples in two categories for $n \in N$ when $N = \{15, 25\}$ and $m = \lfloor \frac{n}{2} \rfloor$. For each category, the coordinate of the stations and terminals are randomly selected from $[0, 10] \times [0, 10]$ with uniform distribution and also the potential of the stations is a integer random number that is Selected from interval $[1, 100]$ randomly (See [27,18]). In all implementation, $r_1 = 0.5$ and $r_2 = 2$. In addition, C_1 and C_2 are considered as first and Second maximum $\frac{1}{c_{kl}}$, respectively. We solve the problem for $p = \lfloor \frac{m}{2} \rfloor$. The distance (in Kilometers) traveled by buses of terminal k for allocating to station l, a random number from interval $[50, 100]$ is selected for input. The number of routs with special lines and the number of BRT buses, integer random numbers from interval $[0, 10]$ are selected for outputs. All implementation are done by IBM ILOG CPLEX Optimization Studio 12.6.1 in Visual Studio 2010 environment. The results of our proposed genetic algorithm for both examples are given in tables 1 and 2, respectively. In first column of the tables, the obtained amounts of the maximum efficiency are given that they show relatively high amount of efficiency. From the maximum efficiency of 15 on the case $n = 15$, we have amount of 12.2062 and from the maximum efficiency of 25 on the case $n = 25$, we have amount of 21.7117. But in second column given the maximum Serving amount , the efficiency is less than one. As it is given in third column, allocating $w_1 = 0.8$ and $w_2 = 0.2$, having small decrease on the Serving amount comparing with the optimal amount, the efficiency is significantly increased comparing with second column. But by allocating $w_1 = 0.5$ and $w_2 = 0.5$ in last column, the Serving amount is decreased more comparing with its preceding column and the efficiency is slightly increased. Therefore, by combing these two problems and allocating suitable amounts to the weights, the terminals with relatively high serving amount and efficiency with be built.

Table 1: The results of the genetic algorithm for the case $n = 15$

	Maximum efficiency	Maximum Serving amount	Best solution for$w_1 = 0.8$ and $w_2 = 0.2$	Best solution for$w_1 = 0.5$ and $w_2 = 0.5$
The weight of the BTLP objective function	0	1	0.8	0.5
The weight of the DEA objective function	1	0	0.2	0.5
Number of established centers	3	3	3	3
Number of selected allocations	15	15	15	15
The amount of the solution for DEA	12.2062	0.164	8.54823	8.78532
The amount of the solution for BTLP	137.681	291.535	289.938	287.48
Time	12.107	11.794	15.988	17.391

6 Conclusions and future work

Here, we studied the bus terminal location and data envelopment analysis problems. We presented the combined model of BTLP and DEA with multi-objective approach In addition to having optimal location, we have patterns with relatively high efficiency of the functional aspect. Then, we proposed a

Table 2: The results of the genetic algorithm for the case $n = 25$

	Maximum efficiency	Maximum Serving amount	Best solution for $w_1 = 0.8$ and $w_2 = 0.2$	Best solution for $w_1 = 0.5$ and $w_2 = 0.5$
The weight of the BTLP objective function	0	1	0.8	0.5
The weight of the DEA objective function	1	0	0.2	0.5
Number of established centers	6	6	6	6
Number of selected allocations	25	25	25	25
The amount of the solution for DEA	21.717	0.215	12.4221	14.9191
The amount of the solution for BTLP	286.623	1897.69	1811.18	1302.25
Time	22.171	22.608	25.248	27.975

genetic algorithm to solve the combine model since BTLP is NP-hard. We implemented our proposed algorithm on random examples. According to the numerical results, we concluded that the terminals selected as optimal locations with maximum service desirability do not generally have the maximum efficiency. In addition, locations with maximum efficiency do not generally have the maximum service desirability. But, we achieved a balance between these two objectives by combing the two models.
Note that, if we run our proposed algorithm on examples with large dimension, the execution of the program has stopped in first iterations because of large number of constraints. Future researches could be the execution of the program in new version of CPLEX and change the implementation. We also can propose other meta-heuristic algorithm.

Acknowledgements

The auhors thank the research council of Ferdowsi University of Mashhad and Optimization Laboratory of Ferdowsi University of Mashhad for supporting this work.

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

References

1. Aboolian, R., Berman, O., & Drezner, Z. (2009). The multiple server location problem. Annals Operation Reasearch, 167(1), 337-352.
2. Akin, D., & Kara, D. (2020). Multicriteria analysis of planned intercity bus terminals in the metropolitan city of Istanbul, Turkey. Transportation Research Part A, 132, 465-489.
3. Babaie-Kafaki, S., Ghanbari, R., & Mahdavi-Amiri, N. (2012). An efficient and practically robust hybrid metaheuristic algorithm for solving fuzzy bus terminal location problems. Operational Research, 29, 125.
4. Babaie-Kafaki, S., Ghanbari, R., & Mahdavi-Amiri, N. (2016). Hybridizations of genetic algorithms and neighborhood search metaheuristics for fuzzy bus terminal location problems. Applied Soft Computing, 46, 220-229.
5. Banker, R. D., & Thrall, R. M. (1992). Estimation of returns to seale using data envelopment analysis. European Journal of Operational Research, 62(1), 74-84.

6. Berman, O., & Drezner, Z. (2007). The multiple server location problem. Journal of the Operational Research Society, 58(1), 91-99.
7. Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. European Journal of Operational Research, 2(6), 429-444.
8. Chen, Z.,& Ming, X. (2020). A rough-fuzzy approach integrating best-worst method and data envelopment analysis to multi-criteria selection of smart product service module. Applied Soft Computing, 94.
9. Colli, T., Rao, D. S. P., & Batteseo, G. E. (2005). An introduction to efficiency and productivity analysis. Springer Science & Business Media.
10. Cooper, W. W., Seiford, L. M., & Tone, K. (2006) Introduction to data envelopment analysis and its uses: with DEA-solver software and references. Springer Science & Business Media.
11. Desai, A., & Storbeck, J. E. (1990). A data envelopment analysis for spatial efficiency, Computers. Enviroment and Urban Systems, 14(2), 145-156.
12. Djeni, A., Radojii, N., Mari, M., & Mladenovi, M. (2016). Parallel VNS for bus terminal location problem. Applied Soft Computing, 42, 448-458.
13. Elhajj, M., Ochieng, W. Y. (2020). Urban bus positioning: Location based services and high level system architecture. Case Studies on Transport Policy, 8, 12-21.
14. Eiselt, H. A., & Marianov, V. (2011). Foundations of location analysis. Springer US.
15. Farrell, M. J. (1957). The measurement of productive efficiency. Journal of the Royal Statistic Society, 120, 253-281.
16. Fayyazi, A., Jafari Eskandari, M., mirzaei poueinak, M., & motamedi nia, P. (2014). Solving urban bus terminal location problem using the meta-heuristic cuckoo optimization algorithm. Basic Sciences Applied Research, 3, 4653.
17. Fogarty, T. C. (1989). Varying the probability of mutation in the genetic algorithm. In Proceedings of the 3rd international conference on genetic algorithms (pp. 104-109). Morgan Kaufmann Publishers Inc.
18. Ghanbari, R., & Mahdavi-Amiri, N. (2011). Solving bus terminal location problem using evolutionary algorithms. Applied Soft Computing, 11, 991-999.
19. Holmgren, J. (2013). The efficiency of public transport operations-An evaluation using stochastic frontier analysis. Research in Transportation Economics, 39(1), 50-57.
20. Klimberg, R. k., $ Ratick, S. J. (2008). Modeling data envelopment analysis (DEA) efficient location/allocation decisions. Computers and Operations Research, 35, 457-474.
21. Knowles, J. D., & Corne, D. W. (2000). M-PAES: A memetic algorithm for multiobjective optimization. In Proceedings of the 2000 Congress on Evolutionary Computation. CEC00 (Cat. No. 00TH8512) (Vol. 1, pp. 325-332). IEEE.
22. Krarup, J., & Pruzan, M. (1990). Ingredients of location analysis, in: P.B. Mirchandani. R.L.Francis (Eds.), Discrete Location Theory, Wiley-Interscience, New York, 154.
23. Lao, Y., & Liu, L. (2009). Performance evaluation of bus lines with data envelopment analysis and geographic information systems. Computers, Environment and Urban Systems, 33(4), 247-255.
24. Mecke, S., Schobel, A.,& Wegner, D. (2006). Stop location-complexity and approximation. In 5th Workshop on Algorithmic Methods and Models for Optimization of Railways, Number 06901 in Dagstuhl Seminar proceedings.
25. Miettinen, K. (2012). Nonlinear multiobjective optimization.Vol. 12. Springer Science & Business Media.
26. Munuzuri, J., & Munoz-Diaz, M. L. (2019). Use of DEA to identify urban geographical zones with special difficulty for freight deliveries. Vol. 79. Journal of Transport Geography.
27. Rahdar, S. (2016). A new model for the bus terminal location problem and solving it using two meta-heuristic algorithms. Master Dissertation, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad (in persian).
28. Ruiz-Torres, A. J., & Lpez, F. J. (2004). Using the FDH formulation of DEA to evaluate a multi-criteria problem in parallel machine scheduling. Computers & Industrial Engineering, 47, 107-121.
29. Sabouhi, F., Bozorgi-Amiri, A., Moshref-Javadi, M., & Heydari, M. (2019). An integrated routing and scheduling model for evacuation and commodity distribution in

large-scale disaster relief operations: a case study. Annals of Operations Research, 283(1-2), 643-677.

30. Sinriech, D., & Samakh, E. (1999). A genetic approach to the pick up/delivery station location problem in segmented flow based material handling systems. Journal of manufacturing systems, 18, 81-99.
31. Saeedi, H., Behdani, B., Wiegmans, B., & Zuidwijk, R. (2019). Assessing the technical efficiency of intermodal freight transport chains using a modified network DEA approach. Transportation Research Part E: Logistics and Transportation Review, 126, 66-86.
32. Toloo, M., Sohrabi, B., & Nalchigar, S. (2009). A new method for ranking discovered rules from data mining by DEA. Expert Systems with Applications, 36, 8503-8508.
33. Viton, P. A. (1997). Technical efficiency in multi-mode bus transit: A production frontier analysis. Transportation Research Part B: Methodological, 31, 23-39.
34. Zhang, L., Wang, Y. P., Sun, J., & Yu, B. (2019). The sightseeing bus schedule optimization under Park and Ride System in tourist attractions. Annals of Operations Research, 273(1-2), 587-605.