Scientia Iranica E (2021) 28(4), 2361–2373

SCIENTIA IRANICA

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
http://scientiairanica.sharif.edu

Developing an iterative procedure to estimate origin-destination matrix based on two-point license plate tracking system

M. Teimouri[a], M. Sheikhmohammady[b,*], A. Husseinzadeh Kashan[b], and A.A. Shojaie[a]

a. *Department of Industrial Engineering, Branch of Tehran South, Islamic Azad University, Tehran, Iran.*
b. *Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran.*

Received 12 December 2018; received in revised form 7 August 2019; accepted 20 October 2019

KEYWORDS
Intercity transportation;
Iterative procedures;
Mathematical programming;
Origin-destination matrix;
Road camera detection systems.

Abstract. Origin-Destination (O-D) matrix, one of the most important elements in transportation planning, is usually estimated by various techniques such as mathematical modeling, statistical methods, and heuristic approaches. Since usage of electronic devices is rapidly increasing which helps decision-makers improve models' capabilities, an iterative procedure was proposed in this paper to estimate the O-D matrix through the detection of vehicles' license plates. The main concept is to track vehicles on the first and last links equipped with plate camera over the shortest path from origins to destinations. A two-step procedure and mathematical models were developed to adjust assigning the passing traffic to the network links by minimizing deviations between the observed and estimated truck traffic volumes. The proposed procedure was elaborated using an illustrative example and it was validated using experimental road network that covered seven eastern provinces of Iran including 310 nodes, 400 two-way edges, and around 3600 origin and destination pairs. Results revealed that the proposed procedure could estimate O-D matrix when the network links were optimally located and equipped with road camera detection systems. In addition, similar to other heuristic approaches, the proposed procedure is sensitive to the number of iterations on the estimation accuracy.

1. Introduction

1.1. Origin-destination matrix estimation

Estimation of the Origin-Destination matrix (O-D for short) is known as one of the most important issues in transport planning and traffic engineering. This matrix indicates the distribution of traffic or trips between trip generation and trip attraction areas over a transport network and eventually, estimates traffic flow over the network links [1]. The area under transport study is divided into "n" zones followed by a matrix that is commonly depicted on two sides of origin and destination notated respectively by i and j and the number of trips as T_{ij}. For instance, $T_{ij} = 1200$ implies that the number of trips from zone i to zone j is estimated to be 1200. In general, it is possible to directly estimate O-D matrix elements by filling out questionnaires as well as directly utilizing the methods such as mathematical or allocation models based on specific transport and traffic parameters. The direct estimation method has been rarely considered because it is time-consuming

*. *Corresponding author. Tel.: +98 21 82884394;*
E-mail addresses: Teimoori54@gmail.com (M. Teimouri); msheikhm@modares.ac.ir (M. Sheikhmohammady); a.kashan@modares.ac.ir (A. Husseinzadeh Kashan); a.shojaie@azad.ac.ir (A.A. Shojaie)

doi: 10.24200/sci.2019.52460.2726

and costly, while it must be necessarily updated after a few years. In contrast, indirect estimation methods, which are repeatedly observed on the practical studies, need less data, budget, and time. One subset of these methods is developed based on traffic volume in the network arcs (edges or links) where demand for couples of O-D is estimated by observing the flow of the selected arcs [2]. The main objective of indirect estimation methods is to distribute traffic volumes over the network and to make adjustments based on what have been observed. In such cases, the O-D matrix is mainly obtained according to the network structure and the amount of traffic flow. Therefore, the quality of estimated O-D matrix depends on the accuracy of input data and the location of traffic counters [3]. A review of the literature shows that the quality of the estimated O-D matrix depends on many factors including assumptions and methods of traffic allocation, the quality of data collected from traffic counters, structure of network links, and eventually locations and the numbers of traffic counters. The last factor is the most important one because the number of traffic counters is usually limited by available funds and other constraints [4]. One of the most important studies to determine the location of traffic counters was done by Yang and Zhou [5]. They studied clues as the source of many kinds or researches. It is not possible to cover all arcs due to financial constraints. Therefore, a software package was used for determining traffic counters as in the study of Ehlert et al. [6], in which covering more important O-D pairs was desirable to directly influence the allocation of traffic counters.

1.2. O-D estimation modeling

A literature review shows that many studies have been conducted to estimate the O-D matrix using indirect methods. A systematic method was proposed to determine the number and the optimal arcs to be counted which ultimately lead to the definition of rules relevant to the selection of optimal traffic counters. These rules capture the following conditions:

1. Covering all O-D couples;
2. Maximum flow ratio;
3. Maximum flow counting;
4. Arc independency [5,7].

Although future researches such as Larsson et al. [8], Cipriani et al. [9], and Yang et al. [3] will introduce new laws besides the above rules, Wang et al. [4] formulated and applied them to hypothetical networks of different dimensions and proved that none of the above rules could be overcome by other rules for all networks and scenarios. In order to apply the above rules, determining the location of devices is another problem. The problem of detecting devices location was developed using computer simulation programs to determine the upper and lower limits of sensors [10] to count and cover the maximum traffic. Results revealed that the application of appropriate models with fewer sensors could reach the limit of traffic counted devices in the estimation of the O-D matrix.

In practice, many rules and techniques have been proposed for determining the optimal number and locations of counter devices based on old O-D matrix and financial constraints to maximize the coverage of routes using the mathematical model [11]. Development of two-level optimization models has also been observed in the literature such as colony optimization where the model maximizes the coverage of the number of O-D pairs at the first level and minimizes the number of traffic counting stations at the second level [12]. Traffic counters are located to estimate the O-D matrix by selecting optimal paths using meta-heuristic methods [13]. Uncertainty is another concern, about which Fei et al. [14] expanded their previous studies to determine the location of traffic counters in a network by reducing uncertainty in the estimation of O-D matrix and modification of its effects on data collection. In case of limited budget, utilizing methods that improve the accuracy of O-D matrix is common, as can be frequently found in the literature in which the entropy maximization method can even be applied to improve the accuracy of O-D estimation [15]. Using the data captured by the license plate systems is another technique to estimate the O-D matrix elements [16]. In this respect, license plate counts between the two cameras were considered as input and then, the models were developed based on the least squared error and converted into a linear model, seen as part of the O-D matrix. They confirmed that there was an estimation method that could examine daily variations of O-D patterns in the highway network. This technique was later improved by portable license plate readers [17]. Much improvement was observed because of considering probability while dynamic movement was studied in the region of Riga by Savrasovs and Pticina [18] to estimate the probability of detecting destinations by recording vehicle type and plate number after dividing traffic into proper routes.

Network traffic density and uncertainty in traffic flow are important. Yang et al. [3] proposed a two-step algorithm for estimating the O-D matrix in which the selection of paths and their dispersion parameters using partial traffic counters is recommended to be applied in a dense network. They developed a nonlinear optimization model that included a dynamic dispersion parameter followed by calculating the minimal squares of total matrix errors and repetition of service matching models to achieve a better convergence. Performance evaluation in a hypothetical network was conducted using input data, while the implementation and testing of a wide range of various coefficients were repeated followed by estimating the mean square error at each stage for O-D and link demand [3]. Hu et al. [19] changed the pattern of estimating O-D matrix by solving the problem of locating traffic counter points on the network. The objective was to determine the minimum number of traffic counting arcs in the network resulting from those arcs that had a basic role in the arc-path matrix-vector. Statistical methods were utilized for estimating the O-D matrix based on numerical data. Static and dynamic estimations of the O-D matrix, estimation of matrix reliability, determination of a set of traffic stations, the number of links needed to obtain maximum information are the main issues to estimate a reliable O-D matrix in this field [1].

1.3. Toward new technologies

New technologies applied to installation cameras and traffic control sensors help determine the optimum arcs for flow counting and, eventually, estimate the O-D matrix. Given the advances made in the field of intelligent transport systems technology and various measuring instruments and observation, the problem of locating traffic counters on the network has found new formulations. For instance, in recent years, plate registration and auto-recognition techniques have been widely used to locate traffic counters on the network [20]. Increasing the number of traffic counter stations or other vehicles equipment such as vehicle license plates practically enhances the precision of O-D matrix estimation. However, the resource constraint problem makes the above-mentioned methods not always feasible. In this respect, it is essential to achieve the optimal number of traffic counter stations and their position in the network [21].

Mobile phones are increasingly used for gathering data on traveling cars over the network to estimate the O-D matrix. Iqbal et al. [22] developed a methodology to utilize a combination of traffic counters and mobile phone data for estimating O-D matrix which checks the traveling cars between mobile towers at time window intervals. They revealed that using mobile phone data was more economical than traditional survey methods developed based on traffic counters. A combination of route choice and traffic counter data was also utilized to estimate O-D matrix in congested networks [23] where the basic principles of stochastic user equilibrium were assumed for route selection. Other studies such as Alexander et al. [24] in this field were conducted to gain a more accurate estimation of trip purposes including home and work as well as other purposes based on movement time and trip frequency. In public transportation, checking smart card data is known as a good solution to improve the accuracy of O-D matrix estimation for passengers without private cars and non-walking trips [25].

1.4. Vision

Literature review indicates that researches on estimation of O-D matrix using traffic counters data include the operational status and practical techniques such as mathematical modeling, simulation, and statistical methods; however, license plate recognition cameras and other Information Technology (IT) devices have been used in recent years. Methods usually function based on the traffic volume of all vehicles detected by their plates over the network, but the fundamental issues are extracted from two different perspectives compared to the previous studies. The first is to consider the concept of vehicle tracking using license plate detection cameras for those vehicles across the inter-city network passing through the selected paths. The second is to estimate the O-D matrix for freight transportation by just one or two detection points; in the present research work, even trucks may be detected by more cameras. So, a two-stage mathematical model is presented. The routes for passage of vehicle fleets on the wide network are determined at the first stage, and license plate detection equipment is used for predicting origins and destinations. The second part is designed as an iterative procedure which takes the first solution to the best position gradually by minimizing the deviation between observed and estimated traffic volumes on the network links. In other words, the main purpose of this paper is to estimate the O-D demand in the freight transport network using the data derived from tracking vehicle fleets by plate recognition cameras located in an intercity road network. This innovation makes it possible for transportation professionals to reduce the estimation errors of O-D matrix.

This article is organized into five main sections. After introduction where topics and relevant studies are discussed, the proposed iterative procedure and the developed mathematical model are explained in the second section in detail. To explain how the procedure works, an illustrative example is then discussed in the third section, followed by more discussion about the case study as well as experimental data together with numerical analysis in the fourth section. A brief summary of the research work and recommendations for further studies are finally discussed in the last section.

2. Iterative procedure and mathematical model

In this research work, an iterative procedure is proposed to solve the problem of estimating O-D matrix. So, this section is divided into two parts. First, the overall view of the proposed iterative procedure is elaborated, followed by developing mathematical models in the second part. The main concept behind the iterative procedure is the limitation of linear programming which assigns non-zero elements to O-D matrix known as basic variables in mathematical programming [26], in which the above procedure gradually updates them.

2.1. Iterative procedure

The iterative procedure includes a number of stages which will be discussed in the following. At the initial stage, the network specifications including nodes, links, or edges, which connect network nodes, and the distances between nodes are set. The O-D pairs known as candidates are then defined according to transport status. Each O-D pair should include the number of vehicles passing from origin to destination. All the shortest paths from origins and destinations are then determined by defining link involvement for all O-D pairs. Assuming that drivers select the shortest paths over the network, each link may be assigned to several O-D pairs. Therefore, solving the initial mathematical model (at the first stage) ensures the involvement of all links for all O-D pairs, which are used for tackling the mathematical model at the next stage.

Input data are defined as a unique pattern that has been specifically proposed in this research work. Defining this pattern represents the novelty of this research in which trucks are detected over the network by their unique plaque numbers. In this case, a short interval time is considered to pass consequent links followed by detecting the vehicles on other links. The first and the last detection times are investigated based on the time interval of each truck. The number of vehicles is defined and observed for the first time on the first link and for the last time on the last link. Therefore, the structure of input data is a four-dimension table including the first and the last passing links, each of which is separately defined by two start and end nodes. In this case, the objective function is to minimize the difference between the observed and estimated numbers of vehicles throughout the equipped links. After running the second stage, the first estimation of the O-D matrix components is obtained. They will be used as input data for the iterative procedure. From now on, the iterative procedure will be applied by updating variables at the upper and lower limits for the estimated O-D matrix elements until no significant changes are observed in two continuous iterations. The root of the average square errors of the assigned and observed transport demands for all O-D pairs is used as the stopping criterion. Minimizing the difference between the assigned and observed traffic volume over the equipped links is considered as the objective function for all iterations. Figure 1 depicts the overall view of the proposed procedure through which iterative steps are depicted by thick arrows.

2.2. Developing mathematical model

- **Basic Concept.** The basic concept used in developing the mathematical model to estimate the O-D matrix lies in tracking and detecting the movement of trucks on the network arcs. According to Figure 2, it is assumed that the vehicles moving from the

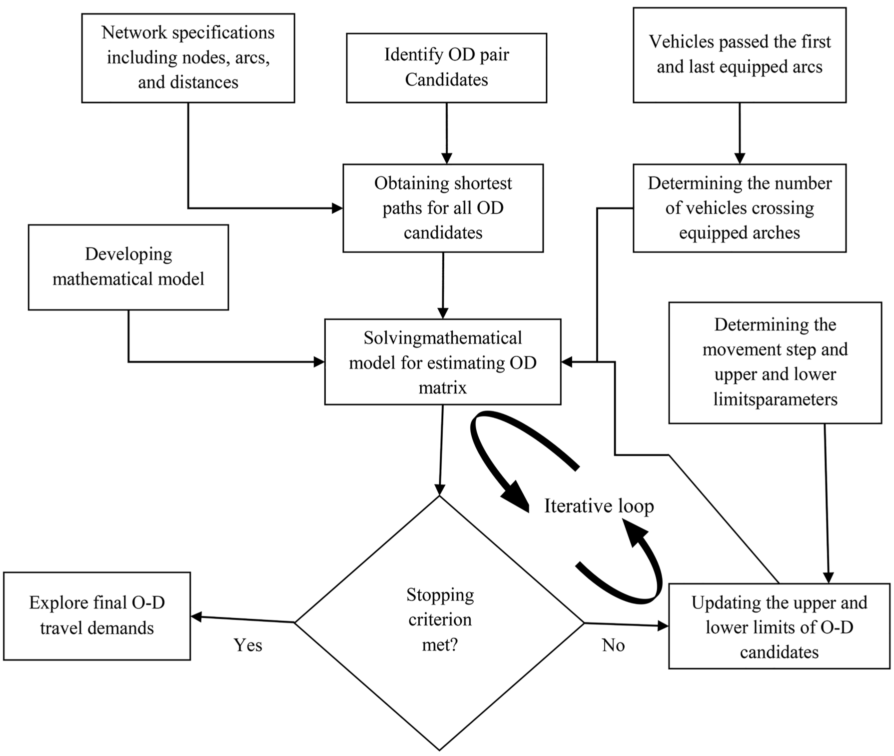

Figure 1. Overall view of the proposed procedure.

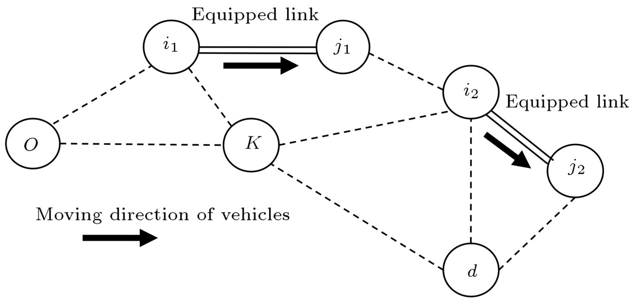

Figure 2. The general pattern of moving vehicles detected over the equipped links.

origin node "o" to the destination node "d" are firstly detected on arc $(i_1 - j_1)$ and detected on arc $(i_2 - j_2)$ for the last time, where the direction of the vehicle movement is also depicted by an arrow over equipped links. It is assumed that the location of equipped links is already known; therefore, locating arcs for installing the license plate system is out of the scope of this research work.

- **Notations and parameters.** Parameters "i" and "j" are node identifiers. Symbol "G" is the road network composed of nodes and their corresponding arcs. Each arc is represented by the symbol (i, j) where "i" and "j" are the start and end nodes, respectively. In the intercity road networks, roads are commonly bidirectional; therefore, the return arc is defined by similar properties defined in Eq. (1):

$$(i, j) \quad \text{and} \quad (j, i) \in G. \tag{1}$$

L_{ij} is the length of arc (i, j) or distance from node "i" to node "j". For the return arc, the length is equal to the main one as defined by Eq. (2):

$$L_{ij} = L_{ji} \in G. \tag{2}$$

CN (o,d) is the set of candidate O-D pairs. They may be old O-D pairs, but it is possible to add more pairs. Following the basic concept, the number of vehicle fleets is known and detected through equipped arcs, defined as the following parameters.

RP_{i1j1}^{i2j2}: The number of vehicles moving from their origins to destinations detected for the first time on arc $(i_1 - j_1)$ and for the last time on arc $(i_2 - j_2)$.

Because there are many short paths over the network, it is necessary to consider such cases in which trucks are detected by one equipped link through the selected path. Mathematical modeling can satisfy the above concern if equipped links are considered the same. So, to improve the model performance in short-distance O-D pairs, it is possible to consider the detection of vehicles by only one equipped link, as defined in Eq. (3):

$$(i_1, j_1) = (i_2, j_2) \qquad \text{for cases in which vehicles pass one equipped arc.} \tag{3}$$

The number of vehicles that pass through each equipped arc is now defined by FM_{ij}. This number represents the total number of vehicles detected on arc $(i - j)$. Another parameter required to compare the estimated and current O-D matrix elements is defined as follows:

M_{od}: The number of vehicles currently moving from origin "o" to destination "d". This is also known as old O-D matrix.

Given that the above parameters are known, two decision variables are defined. The first variable is to determine the shortest routes over the network and the second variable is to determine the trucks passing through O-D pairs. Since the first part of the model is to determine all routes for all O-D candidates, the first decision variable is defined by Eq. (4) as a binary variable.

$$\begin{cases} X_{ij}^{od} = 1 & \text{If the link } (i,j) \text{ is located on the shortest path from origin O to destination D,} \\ 0 & \text{Otherwise.} \end{cases} \tag{4}$$

The second decision variable that determines the assigned O-D matrix elements is defined as follows:

Y_{od}: The number of vehicles moving from origin "o" to destination "d".

- **Modeling procedure.** It is assumed that traffic volume does not affect vehicle routes over an intercity network. All truck drivers select the shortest paths from origin to destination without congestion. According to the principles of users and system equilibrium assumptions [27], determining the shortest paths for all O-D candidates is the same as the total sum of objective function values for individual paths. Therefore, the objective function at the initial stage can be defined by Eq. (5), where Z_1 is the sum of the total traveled distances between origins and destinations. This equation guarantees that all drivers select their shortest routes from their origins to the corresponding destinations.

$$\min Z_1 = \sum_{(o,d)\in CN} \sum_{(i,j)\in G} L_{ij} \times X_{ij}^{od}. \tag{5}$$

The constraint existing at the initial stage is to keep the route from origin to destination seamless. This constraint is formulated by Eq. (6) where $E_x(j)$ is the set of arcs exiting from "j" and $E_n(j)$ is the set of arcs entering node "j". More details about general modeling are available in [26] and about its practical usage in [28].

$$\sum_{i\in E_x(j)} X_{ij}^{od} - \sum_{i\in E_n(j)} X_{ij}^{od} = \begin{Bmatrix} 1 & \text{if} & j=o \\ -1 & \text{if} & j=d \\ 0 & \text{if} & O.W. \end{Bmatrix}$$

$$\forall j \in G \quad \text{and} \quad (o,d) \in CN. \tag{6}$$

The first priority of the proposed procedure is to determine the links located in specific O-D pairs. Therefore, the result shows if the link (i, j) is located on the path (o, d) or not. From now on, the symbol X_{ij}^{od} will not be considered a variable; rather, it will be used as a parameter because it has been firstly assigned by 0 or 1 for all O-D pairs by the routing model. In the second part, the objective function is to determine the number of vehicles moving from the origin "o" to the destination "d" (O-D candidate pairs). Accordingly, the goal is to minimize the total absolute value of error obtained by the difference between the observed and estimated traffic volumes on equipped arcs defined by Eq. (7). The remarkable point is that Eq. (7) covers the equipped arcs only; therefore, the involved or relevant arcs are represented by Eq. (8). In fact, Eq. (8) guarantees avoiding duplications when trucks are detected by one camera installed over the link (i_1, j_1), which is also recounted by Eq. (3). In this case, when a link is counted twice, Eq. (8) deletes the redundant truck counting.

$$\min Z_2 = \sum_{(i,j)\in G} \left| \sum_{(o,d)\in CN} X_{ij}^{od} \times Y_{od} - FM_{ij} \right|, \tag{7}$$

$$(i,j) = (i_1, j_1) \cup (i_2, j_2) \in G. \tag{8}$$

Eq. (7) is formulated as an absolute value, which makes it a non-linear structure. In order to convert it to a linear equation, positive and negative deviations are defined. In this case, Eq. (7) is replaced by Eq. (9) into Eq. (10), where PS_{ij} and NG_{ij} are positive and negative deviations of the estimated and the observed traffic volumes for arc $(i - j)$, respectively.

$$\min Z_2 = \sum_{(i,j)\in G} PS_{ij} + NG_{ij}, \tag{9}$$

$$\sum_{(o.d)\in CN} X_{ij}^{od} \times Y_{od} - FM_{ij} - PS_{ij} + NG_{ij} = 0$$

$$\forall (i,j) \in G. \tag{10}$$

In addition to the above-mentioned constraints, the total vehicles assigned to each equipped arc should match the total detected vehicles. These constraints are formulated by Eqs. (11) and (12). Eq. (11) satisfies the number of vehicles that pass through the first equipped arc and Eq. (12) satisfies the number of vehicle fleets in the last equipped arc for each O-D candidate pair.

$$\sum_{(o,d)\in CN} X_{i_1j_1}^{od} \times Y_{od} = \sum_{(i_2,j_2)\in G} RP_{i_1j_1}^{i_2j_2}$$

$$\forall (i_1, j_1) \in G, \qquad (o,d) \in CN, \tag{11}$$

$$\sum_{(o,d)\in CN} X_{i_2j_2}^{od} \times Y_{od} = \sum_{(i_1,j_1)\in G} RP_{i_1j_1}^{i_2j_2}$$

$$\forall (i_2, j_2) \in G, \qquad (o,d) \in CN. \tag{12}$$

2.3. Iterative constraints

In linear programming, the number of basic (non-zero) variables is equal to or less than the number of constraints [26]. In the proposed mathematical models, the number of constraints is equal to that of equipped arcs which is multiplied by 3, as can be seen in Eqs. (10)–(12). In this case, the number of non-zero variables, called basic variables, is less than the number of constraints. Therefore, the number of O-D pairs is larger, which is the reason why the mathematical model does not assign elements for all O-D pairs. In order to solve this problem, an iterative procedure is developed. The proposed procedure turns the traveled assignments to specific values step by step. These specific values are assumed as the values of old O-D pairs defined as the parameter FM_{ij} in the previous sub-section. In the iterative procedure, the obtained results are compared to the previous iteration results. Following a comparison of the results, two cases may occur:

Case 1: After running the mathematical model, the traffic volume assigned to each specific link is less than old O-D or zero. In this case, a lower bound is necessary to restrict the assigned traffic volume for the link. Figure 3 depicts the first situation in which the assigned traffic volume is less than the O-D assigned value used for validation. The assigned values may represent the old O-D pairs. The horizontal axes show different O-D values investigated in the solving procedure. In addition, an upper bound is

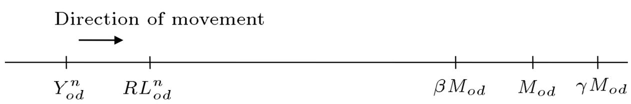

Figure 3. Movement direction when the assigned Origin-Destination (O-D) is less than the old value.

required to restrict the assigned traffic volume since the mathematical model may obtain unusual traffic volume. Based on the above discussion, the urgency to find a better solution is satisfied by Eq. (13), where α is a coefficient for movement speed, β is an acceptable coefficient to increase the lower bound, Y^n_{od} is the traffic volume assigned to each O-D pair at iteration "n", RL^n_{od} is the amount of lower bound, and eventually γ is an acceptable coefficient to raise the upper bound of assigned traffic volume to each O-D pair. The next iteration is conducted to update all the above variables by Eqs. (14) and (15).

If $\quad Y^n_{od} \leq M_{od};$

$$RL^n_{od} = \min\left[Y^n_{od} + \alpha(M_{od} - Y^n_{od}), \quad \beta M_{od}\right], \qquad (13)$$

$$Y^{n+1}_{od} \geq RL^n_{od} \quad \forall od \in OD \quad \text{and} \quad Y^n_{od}\langle M_{od}, \qquad (14)$$

$$Y^{n+1}_{od} \leq \gamma\ M_{od} \quad \forall od \in OD \quad \text{and} \quad Y^n_{od}\langle M_{od}. \qquad (15)$$

Case 2: After running the mathematical model, the assigned traffic volume becomes much greater than the old O-D. In this case, an upper bound is necessary to limit assigned traffic volume to a specific link. Figure 4 depicts the first state in which the assigned traffic volume is greater than the desired condition. In addition, a lower bound is also necessary to restrict the assigned traffic volume since the mathematical model may obtain unusual traffic volume. The urgency to find a better solution is satisfied by Eq. (16), in which all coefficients and variables are the same as those in Case 1 and RU^n_{od} is the upper bound value for the assigned traffic volume of O-D pairs. The next iteration is done to update all the above variables by Eqs. (17) and (18):

If $\quad Y^n_{od} \geq M_{od};$

$$RU^n_{od} = \max\left[Y^n_{od} + \alpha(Y^n_{od} - M_{od}), \quad \gamma M_{od}\right], \qquad (16)$$

$$Y^{n+1}_{od} \leq RU^n_{od} \quad \forall od \in \ OD \quad \text{and} \quad Y^n_{od} \geq M_{od}, \quad (17)$$

$$Y^{n+1}_{od} \geq \beta M_{od} \quad \forall od \in OD \quad \text{and} \quad Y^n_{od} \geq M_{od}. \quad (18)$$

3. Illustrative example

In order to better understand how the proposed model works, an illustrative example is discussed. It is

Figure 4. Movement direction when the assigned Origin-Destination (O-D) is greater than the old value.

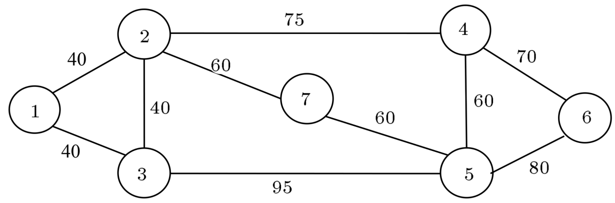

Figure 5. Illustrative network structure.

assumed that there is a simple network shown in Figure 5. The network consists of 7 nodes and 10 double-sided arcs (20 links). The numbers located on the arcs represent the lengths and the connections are depicted as single lines, which imply that there are two directions available. In addition, nine O-D pairs are considered, as tabulated Table 1.

Travel demand and the shortest path are given in Table 1, which contains the length of the shortest path for each O-D pair. For example, the first row shows data for O-D pairs from Nodes 1 to 5. The current travel O-D demand includes 500 vehicles and the shortest route from Origin 1 to Destination 5 is 1-3-5 which involves two links (1,3) and (3,5). According to Figure 5, the length of the path is 135 km, as shown in the last column. Other rows show more information for the remaining O-D pairs. It is assumed that Arcs (1,3), (2,4), (2,7), (3,2), (3,5), (2,4), (6,5), and (7,5) are equipped with license plate reader systems. The vehicles moving from the origin "o" to the destination "d" were detected for the first time in the arc (i_1, j_1) and for the last time in the arc (i_2, j_2). All travel data on vehicles passing arcs are tabulated in Table 2. For example, the first row indicates that there are 500 vehicles that move from a specified origin to its destination for the first time in the arc (1,3) and for the last time in the arc (3,5). The other rows similarly illustrate a larger number of vehicles for the other links.

To estimate the O-D matrix elements, the proposed procedure and the mathematical model are utilized by using the well-known software of GAMS where the moving coefficient "α" is set to 0.05, β is set to 0.75, and γ is set to 1.25. The results are represented in Table 3.

According to Table 3, at the initial stage ($n = 0$), the assigned O-D pairs are less than the old or the desired demand in some O-D pairs; however, some are zero and some are greater than the old demands. For example, in the third row, the current demand for the O-D pairs 3 to 4 is 200 in number, but it is assigned zero at the initial stage. The root Mean Absolute Error (MAE) is also calculated as 211. As the procedure evolves, the assigned O-D pairs move gradually to a steady state where the assigned O-D pairs for (3 to 4) changed to 108 at the third iteration. Running the proposed procedure improves the criterion which

Table 1. Data assumed for the illustrative example.

Row	Origin	Destination	O-D demand	Shortest path	Length
1	1	5	500	$1 \to 3 \to 5$	135
2	1	6	650	$1 \to 2 \to 4 \to 6$	185
3	3	4	200	$3 \to 2 \to 4$	115
4	3	7	350	$3 \to 2 \to 7$	100
5	5	2	300	$5 \to 3 \to 2$	135
6	6	2	450	$6 \to 4 \to 2$	145
7	6	5	400	$6 \to 5$	80
8	6	7	450	$6 \to 5 \to 7$	140
9	7	6	250	$7 \to 5 \to 6$	140

Table 2. Equipped links involved in passing vehicles (illustrative example)

Row	The first arc that records the plaque	The last arc that registers the plaque	Number of vehicles	Row	The first arc that records the plaque	The last arc that registers the plaque	Number of vehicles
1	(1,3)	(3,5)	500	5	(3,2)	(3,2)	300
2	(2,4)	(2,4)	850	6	(4,2)	(4,2)	450
3	(3,2)	(2,4)	200	7	(6,5)	(6,5)	850
4	(3,2)	(2,7)	350	8	(7,5)	(7,5)	250

Table 3. Origin-Destination (O-D) pairs estimation matrix in different number of iterations (illustrative example).

Row	Origin	Destination	Existing demand	Iteration 0	Iteration 3	Iteration 5	Iteration 10	Iteration 20
1	1	5	500	200	500	500	500	500
2	1	6	650	1050	942	886	850	850
3	3	4	200	0	108	164	200	200
4	3	7	350	350	350	350	350	350
5	5	2	300	500	392	336	300	300
6	6	2	450	450	450	450	450	450
7	6	5	400	0	400	400	400	400
8	6	7	450	850	450	450	450	450
9	7	6	250	250	250	250	250	250
MAE*				211	53	34	22	22

* MAE: Mean Absolute Error.

expresses that all the assigned O-D pairs should reach a steady state at the 10th iteration. No more iteration is required because no change is made after Iteration 10. Therefore, the final O-D pairs are now estimated at Iteration 10. In order to check the convergence rate of the procedure, MAE is used to calculate the difference between the observed and assigned trucks for each equipped link. The last row of Table 3 shows the smooth decreasing pattern on MAE, while the number of iterations increases.

4. Experimental analysis

4.1. Case study

In order to utilize the proposed procedure and the developed mathematical model to resolve real-world problems, an intercity road network was selected in the eastern part of Iran including seven provinces of Golestan, North Khorasan, Khorasan Razavi, South Khorasan, Sistan and Baluchestan, Kerman, and Semnan colored in yellow in Figure 6. The land size of these

Figure 6. An overall view of the case study highlighted in yellow.

provinces constitutes around 47% of the country where the total intercity road length is 18,497 kilometers, around 20% of the total road length. The experimental network consists of 310 nodes, 400 one-way edges (or 800 two-way edges), and 3609 O-D pairs extracted from the issued transport documents. Data for O-D pairs were gathered for one year (March 21, 2016 to March 20, 2017). The network structure is graphed manually to make adjustments for running the model. To improve data quality, border points are defined as origin or destination nodes. For example, total demand for transport from origin to destination for the provinces of Khorasan Razavi, North Khorasan, South Khorasan, and Semnan to western provinces of Iran is concentrated in Garmsar, i.e., the west-northern node of the case study. This process is similarly carried out for all O-D pairs and eight border points are defined as dummy nodes. Therefore, each dummy node represents the sum of transport supply or demand, defined by the number of trucks, for all O-Ds entering or departing from the area under study. More explanations on how to define boarder nodes are available in [28] since the same pattern is also used in this research work.

4.2. Procedure implementation

Despite the existing large-scale network, the well-known optimization software of GAMS, algorithm CPLEX, is able to solve the mathematical model. Hence, no heuristic approach to solving this problem is necessary. Another coding pattern is also used for nodes numbered sequentially from 1001 to 1310. Since detecting cameras are usually used for other purposes such as speed enforcement, there are 145 cameras installed over the network. The O-D pairs with more than 50 trips in a year were distributed over the network. The results are summarized in Table 4. In this table, the first three rows and the last three rows are shown. The first and the last equipped links are also identified in the second and third columns, respectively.

The application of the proposed procedure obtains the estimated O-D matrix over the network at

Table 4. The number of vehicles passing over the first and the last links.

Row	First detecting link	Last detecting link	Number of vehicles
1	(1179, 1178)	(1307, 1001)	13965453
2	(1196, 1194)	(1141, 1121)	4842134
3	(1196, 1194)	(1235, 1131)	2417879
...	...	...	...
...	...	...	...
2509	(1080, 1089)	(1025, 1053)	51
2510	(1035, 1171)	(1177, 1178)	51
2511	(1022, 1047)	(1081, 1079)	51

different iterations, as shown in Table 5. In addition to the observed and assigned O-D values, the Root Mean Square Error (RMSE) is also calculated. To compare the model results and the observed trucks assigned to the links, the well-known criterion of RMSE is investigated. The difference between the observed and assigned trucks is calculated. This criterion shows how the model converges to the real situation. According to Table 5, O-D pairs converge to the specific values after the 75th iteration. It means that the proposed procedure is capable to estimate O-D pairs after 75 iterations.

Investigation of the RMSE that the convergence speed for the initial steps is higher than that for the final steps, as shown in Figure 7. It is implied that the procedure can be used in a few steps if the network is large sized and the accuracy is not strongly important.

4.3. Sensitivity analysis

As mentioned in Section 2.3, three scalars are used for converging assigned O-D pairs to the final solutions. Sensitivity analysis should be carried out based on the

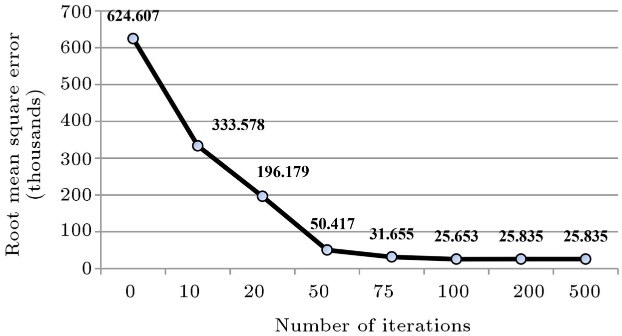

Figure 7. Root mean square errors based on the number of iterations.

performances of the parameters. To check the performance, different ways are proposed and examined. For this purpose, the upper and lower limits are considered in a range of responses as values of β and γ for different values of 0.75, 0.85, and 0.95, respectively, with the corresponding limit values being above 1.25, 1.15, and 1.05. The above coefficients show that the upper and lower bounds have the same interval being adjusted to 1. However, it is not necessary to do so and researchers may use different intervals for this matter. Moving speed is also checked by values of 0.01, 0.02, and 0.05. The coefficient $\beta = 0.75$ implies that the lower bound for the estimated O-D element is not allowed to be less than 75% of the old O-D. The upper bound is now restricted up to 125% of current O-D, where $\gamma = 1.25$. Steps are now forced by setting the parameter α, which justifies the speed of convergence to the upper or lower bounds. Results summarized in Table 6 reveal the following concluding remarks:

- The error of the proposed model was reduced upon adding more iterations;
- O-D pairs reached a steady state after 200 iterations,

Table 5. Estimated Origin-Destination (O-D) matrix at different iterations.

Row	O	D	O-D	Iteration 0	Iteration 10	Iteration 20	Iteration 50	Iteration 75	Iteration 100	Iteration 200	Iteration 500
1	1001	1002	11482	23079	18426	15639	14353	14353	14353	14353	14353
2	1001	1004	48528	54924	52198	50543	49869	49864	49879	49884	49884
3	1001	1007	1076727	0	432051	693502	807545	807545	807545	807545	807545
...	...	...	...	...	...	...	...	...	...	...	...
1804	1126	1025	786	0	315	504	590	590	590	590	590
1805	1126	1027	2655	0	1065	1703	1991	1991	1991	1991	1991
1806	1126	1028	35	0	14	22	26	26	26	26	26
...	...	...	...	...	...	...	...	...	...	...	...
3607	1281	1131	126	0	95	95	158	158	158	158	158
3608	1281	1152	6	0	2	4	8	8	8	8	8
3609	1281	1159	17	0	7	11	13	13	13	13	13
RMSE				624607	333578	196179	50417	31655	25653	25835	25835

Table 6. Sensitivity analysis results using different parameters.

Row	β	γ	α	Iteration 0	Iteration 10	Iteration 20	Iteration 50	Iteration 100	Iteration 200	Iteration 500
1	0.75	1.25	0.01	624607	512321	451183	329621	198005	80884	25645
2	–	–	0.02	624607	450695	365945	197302	80105	25732	25435
3	–	–	0.05	624607	333578	196179	50417	25653	25835	25435
4	0.85	1.15	0.01	624607	504263	452086	331343	198810	72552	17399
5	–	–	0.02	624607	456505	370350	198407	72886	20160	17669
6	–	–	0.05	624607	336288	198916	46356	17386	17669	17669
7	0.95	1.05	0.01	624607	515161	453314	331310	197597	67676	7735
8	–	–	0.02	624607	455118	369860	197622	68434	10166	7678
9	–	–	0.05	624607	334878	197899	38061	7694	7678	7679

indicating that analysts should always observe the variation of the results in running steps instead of the number of iterations;

- Using the lower convergence speed results in better outcomes than upper ones;
- Upper and lower limits brought the results closer to the reality.

Hence, it is concluded that the iterative procedure developed in this research enjoys a good performance to estimate O-D matrix using detection of vehicles over the network.

We have discussed the convergence speed that was obtained by the proposed procedure. In order to check the above consideration, the obtained RMSE are given in Figure 8 for $\beta = 0.75$ and $\gamma = 1.25$ using iteration numbers. Accordingly, the procedure finally converges to the same solution, but the coefficient α plays an important role throughout the calculation process.

5. Summary and conclusion

The Origin-Destination (O-D) matrix estimation is one of the most important steps in transport planning. Since electronic devices that detect vehicles' registration plates are increasingly used in intercity transportation, tracking vehicles is a good idea to estimate the O-D matrix. Therefore, an iterative procedure was proposed in this research work to estimate O-D matrix based on vehicle tracking over the intercity network where trucks were tracked by license plate detection systems. After reviewing the literature, a mathematical model was developed based on the concept of tracking vehicles and an iterative procedure was developed to obtain O-D pairs as well. The objective function was to minimize the residuals obtained from the observed and assigned numbers of vehicles over equipped links. The difference between the assigned and the old O-D pairs was defined as a stopping criterion for the iterative procedure. The proposed procedure was utilized using an illustrative example which consists of 7 nodes and 20 arcs to demonstrate what happens during the running procedure and how to check the procedure accuracy.

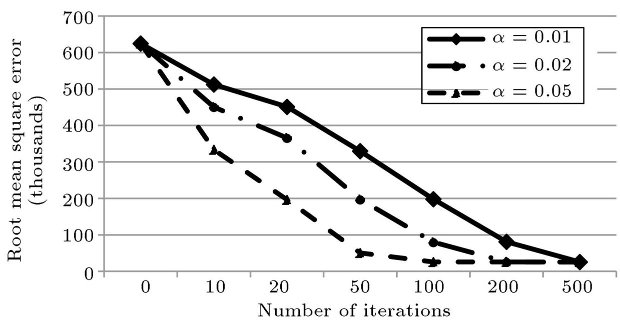

Figure 8. Root mean square errors based on the convergence coefficient of α.

The eastern part of Iran which consists of 310 nodes and 400 two-way edges and 3609 O-D pairs was selected as the case study. After collecting enough data of current O-D matrix over a one-year duration, results were examined for 145 links which were previously equipped for other purposes such as speed management. The results revealed that the proposed procedure, which was developed following the concept of vehicle tracking by two license plate recognition cameras, could accurately estimate the elements of the O-D matrix. Since the installation of electronic devices is widespread, further research is recommended so that researchers would focus on traffic volumes over the links or investigate local or temporary road closure schemes authorized by the respective transport authorities over the network. In such a circumstance, the reliability of link connectivity can be considered to deal with the routing problem, which will be thoroughly explored in the authors' future research work.

References

1. Bera, S. and Rao, K.V. "Estimation of origin-destination matrix from traffic counts: The state of

the art", *European Transport*, **49**, pp. 3–23 (2011).

2. Castillo, E., Menéndez, J.M., and Sánchez-Cambronero, S. "Predicting traffic flow using Bayesian networks", *Transportation Research Part B: Methodological*, **42**(5), pp. 482–509 (2008).
3. Yang, H., Yang, C., and Gan, L. "Models and algorithms for the screen line-based traffic-counting location problems", *Computers & Operations Research*, **33**(3), pp. 836–858 (2006).
4. Wang, N., Gentili, M., and Mirchandani, P. "Model to locate sensors for estimation of static origin-destination volumes given prior flow information", *Transportation Research Record: Journal of the Transportation Research Board*, **2283**, pp. 67–73 (2012).
5. Yang, H. and Zhou, J. "Optimal traffic counting locations for origin-destination matrix estimation", *Transportation Research Part B: Methodological*, **32**(2), pp. 109–126 (1998).
6. Ehlert, A., Bell, M.G., and Grosso, S. "The optimisation of traffic count locations in road networks", *Transportation Research Part B: Methodological*, **40**(6), pp. 460–479 (2006).
7. Yang, H., Iida, Y., and Sasaki, T. "An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts", *Transportation Research Part B: Methodological*, **25**(5), pp. 351–363 (1991).
8. Larsson, T., Lundgren, J.T., and Peterson, A. "Allocation of link flow detectors for origin-destination matrix estimation-A comparative study", *Computer-Aided Civil and Infrastructure Engineering*, **25**(2), pp. 116–131 (2010).
9. Cipriani, E., Fusco, G., Gori, S., et al. "Heuristic methods for the optimal location of road traffic monitoring", *In Intelligent Transportation Systems Conference*, ITSC'06. IEEE, pp. 1072–1077 (2006).
10. Bianco, L., Confessore, G., and Reverberi, P. "A network based model for traffic sensor location with implications on O/D matrix estimates", *Transportation Science*, **35**(1), pp. 50–60 (2001).
11. Mínguez, R., Sánchez-Cambronero, S., Castillo, E., et al. "Optimal traffic plate scanning location for OD trip matrix and route estimation in road networks", *Transportation Research Part B: Methodological*, **44**(2), pp. 282–298 (2010).
12. Sun, D., Chang, Y., and Zhang, L. "An ant colony optimisation model for traffic counting location problem", *In Proceedings of the Institution of Civil Engineers-Transport*, **165**(3), pp. 175–185 (2012).
13. Kim, H.J., Chung, H.I., and Chung, S.Y. "Selection of the optimal traffic counting locations for estimating origin-destination trip matrix", *Journal of the Eastern Asia Society for Transportation Studies*, **5**, pp. 1353–1365 (2003).
14. Fei, X., Mahmassani, H., and Eisenman, S. "Sensor coverage and location for real-time traffic prediction in large-scale networks", *Transportation Research Record: Journal of the Transportation Research Board*, **2039**, pp. 1–15 (2007).
15. Khan, T. and Anderson, M. "Accurately estimating origin/destination matrices in situations with limited traffic counts: Case study huntsville, AL", *International Journal of Traffic and Transportation Engineering*, **5**(3), pp. 64–72 (2016).
16. Asakura, Y., Hato, E., and Kashiwadani, M. "Origin-destination matrices estimation model using automatic vehicle identification data and its application to the Han-Shin expressway network", *Transportation*, **27**(4), pp. 419–438 (2000).
17. Ioli, Edward, D. "Vehicle identification, tracking and enforcement system", U.S. Patent 8,937,559 (2015).
18. Savrasovs, M. and Pticina, I. "Methodology of OD matrix estimation based on video recordings and traffic counts", *Procedia Engineering*, **178**, pp. 289–297 (2017).
19. Hu, S.R., Peeta, S., and Chu, C.H. "Identification of vehicle sensor locations for link-based network traffic applications", *Transportation Research Part B: Methodological*, **43**(8–9), pp. 873–894 (2009).
20. Castillo, E., Gallego, I., Sanchez-Cambronero, S., et al. "Matrix tools for general observability analysis in traffic networks", *IEEE Transactions on Intelligent Transportation Systems*, **11**(4), pp. 799–813 (2010).
21. Yim, P.K. and Lam, W.H. "Evaluation of count location selection methods for estimation of OD matrices", *Journal of Transportation Engineering*, **124**(4), pp. 376–383 (1998).
22. Iqbal, M.S., Choudhury, C.F., Wang, P., et al. "Development of origin-destination matrices using mobile phone call data", *Transportation Research Part C: Emerging Technologies*, **40**, pp. 63–74 (2014).
23. Wang, Y., Ma, X., Liu, Y., et al. "A two-stage algorithm for origin-destination matrices estimation considering dynamic dispersion parameter for route choice", *PloS one*, **11**(1), p. e0146850 (2016).
24. Alexander, L., Jiang, S., Murga, M., et al. "Origin-destination trips by purpose and time of day inferred from mobile phone data", *Transportation Research Part C: Emerging Technologies*, **58**, pp. 240–250 (2015).
25. Alsger, A.A., Mesbah, M., Ferreira, L., et al. "Use of smart card fare data to estimate public transport origin-destination matrix", *Transportation Research Record: Journal of the Transportation Research Board*, **2535**, pp. 88–96 (2015).
26. Taha, H.A., *Operations Research: An Introduction (for VTU)*, Pearson Education India (2005).
27. Castillo, E., Grande, Z., Calviño, A., et al. "A state-of-the-art review of the sensor location, flow observ-

ability, estimation, and prediction problems in traffic networks", *Hindawi Publishing Corporation Journal of Sensors*, **2015**, Article ID 903563, 26 pages (2015). http://dx.doi.org/10.1155/2015/903563

28. Mahmoudabadi, A. and Seyedhosseini, S.M. "Solving hazmat routing problem in chaotic damage severity network under emergency environment", *Transport policy*, **36**, pp. 34–45 (2014).

Biographies

Mohammad Teimouri received his undergraduate degree from Yazd University in 1997. He then received his MSc degree in Socio-Economic Systems Engineering from Tarbiat Modares University Tehran, Iran in 2000. He holds a PhD degree in Industrial Engineering. Dr. Teimouri received his PhD from the Department of Industrial Engineering, Branch of Tehran South, Islamic Azad University. His main research area lies in intelligent transportation systems. Currently, he is the Vice President of Iran's Road Maintenance and Transportation Organization.

Majid Sheikhmohammady received his PhD degree in Systems Design Engineering from University of Waterloo, Canada in January 2009. He was also a Post-Doctoral Fellow at the Department of Systems Design Engineering, University of Waterloo in 2009. He is currently a faculty member at the Department of Industrial and Systems Engineering, Tarbiat Modares University. Since January 2019, he has been an active member of Conflict Analysis Group, Department of Systems Design Engineering, University of Waterloo, Canada. Dr. Sheikhmohammady's primary research interests lie in the analysis of decisions as the intersection of mathematics, engineering, and social sciences. His applications of game theory and related formal techniques include problems in conflict resolution, coalition formation, group decision and negotiation, fair division, and win-win solutions. He has pioneered the development of a methodology to model and analyze multilateral negotiations from a Systems Engineering perspective. More specifically, the new methodology called EMMN predicts the outcome of multilateral negotiations. Generally, Dr. Sheikhmohammady is interested in systems methodologies for innovatively solving real-world complex interdisciplinary problems that lie at the interface of society, technology, and the environment. In addition to his background in Industrial Engineering, he has carried out several research projects in other fields of study including natural resources, international relations, environmental studies, and defense science.

Ali Husseinzadeh Kashan holds degrees in Industrial Engineering from Amirkabir University of Technology (Poly Technique of Tehran), Iran. He worked as a postdoctoral research fellow at the Department of Industrial Engineering and Management Systems with the financial support of Iran National Elite foundations. Dr. Kashan is currently an Associate Professor at the Department of Industrial and Systems Engineering, Tarbiat Modares University and has been active in the applied optimization research field since 2004. His research focuses on modeling and solving hard combinatorial optimization problems in areas such as logistics and supply networks, revenue management and pricing, resource scheduling, grouping problems, financial engineering, etc. Having proposed solution methodologies for real-world engineering design problems, he introduced several intelligent optimization procedures inspired from natural phenomena such as League Championship Algorithm (LCA), Optics Inspired Optimization (OIO), Find-Fix-Finish-Exploit-Analyze (F3EA) metaheuristic algorithm, and Grouping Evolution Strategies (GES). Dr. Kashan has published over 100 peer-reviewed journal and conference papers and has served as a referee for several outstanding journals such as IEEE Transactions on Evolutionary Computations, Omega, Computers & Operations Research, Journal of the Operational Research Society, Computers & Industrial Engineering, International Journal of Production Research, Information Sciences, Applied Soft Computing, Ecological Informatics, Engineering Optimization, Optimal Control and Applications, etc. He has received several awards from Iran National Elite Foundation and in 2016, he was honored by the Academy of Sciences of Iran as the "outstanding young scientist of Industrial Engineering".

Amir Abbas Shojaie holds PhD in Industrial Engineering. He received his PhD from the Department of Sciences and Researches, Islamic Azad University Tehran, Iran in 2003. Currently, he is a faculty member at the Department of Industrial Engineering, Branch of Tehran South, Islamic Azad University. His main research areas include transportation planning, production policies, and control project.