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, A. A. Shojaei 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

Abstract

Origin-Destination 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 using electronic devices is rapidly increased to help decision makers to improve models’ capabilities, an iterative procedure is proposed in this paper to estimate the O-D Matrix according to vehicles’ license plates detection. The main concept is to track vehicles on the first and the last links equipped by plate camera over the shortest path from origins to destinations. A two-step procedure and mathematical models are developed to adjust assigned the passing traffic to the network links by minimizing deviations between the observed and estimated truck traffic volumes. The proposed procedure is explained by an illustrative example followed by validation using experimental road network that covers 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 is capable to estimate O-D matrix when the network links are optimally located and equipped by road camera detection systems. In addition, such as the other heuristic approaches, the proposed procedure is sensitive to the number of iterations on the estimation accuracy.

Keywords: Intercity Transportation, Iterative Procedures, Mathematical Programming, Origin-Destination Matrix, Road Camera Detection Systems

Introduction

1.1. Origin-Destination Matrix Estimation

Estimating 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...

*Corresponding Author. Tel: +98 2182884394

E-mail addresses: msheikhm@modares.ac.ir (M. Sheikhmohammady); Teimoori54@gmail.com (M. Teimouri); a.kashan@modares.ac.ir (A. Husseinzadeh kashan); a.shojaie@azad.ac.ir (A. A. Shojaei).
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 which is commonly depicted in two sides of origin and destination, respectively notated by i and j and the number of trips as $T_{ij}$. For instance, $T_{ij}=1200$ means 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 estimating method has been rarely considered because it is time-consuming 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 fewer data, budget, and time. One subset of these methods are developed based on traffic volume in the network arcs (edges or links) where the demand for couples of O-D is estimated by observing the flow of the selected arcs [2]. The main viewpoint behind the indirect estimation methods is to allocate traffic volumes over the network and to modify them based on what are 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]. Reviewing 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, network links structure 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 the clues as the source of further kinds or researches. It is not possible to cover all arcs due to financial constraints. Therefore, software has been developed for determining the traffic counters such as what done by Ehlert et al [6], in which covering more important O-D pairs was desirable to directly influence on the allocation of traffic counters.

### 1.2. O-D Estimation Modeling

Looking at the literature 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 of and the optimal arcs to be counted which are ultimately led to the definition of rules relevant to the selection of the optimal traffic counters. These rules capture the following conditions: 1) covers all O-D couples, 2) maximum flow ratio, 3) maximum flow counting, and 4) arc independency [5, 7]. Although, the next researches such as Larsson et al [8], Cipriani et al [9] and Yang et al [3]; introduced new laws adding to the above rules, and Wang et al formulated and applied them on hypothetical networks of different dimensions and proved that none of the above rules can be overcome by the other rules for all networks and scenarios. In order to apply the above rules, determining the location of devices was another problem. The problem of detection devices location was developed incorporating of using computer simulation program to determine the upper and lower limits of sensors [10] to count and cover the maximum traffic. Results revealed that using appropriate models with
less number of sensors could reach the limit of the 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 as well as considering the financial constraints to maximize the coverage of routes using mathematical model[11]. Developing two-level optimization models has also been observed in the literature such as colony optimization where 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 through selecting optimal paths by using meta-heuristic methods [13]. Uncertainty is another concern where Fei et al [14] expanded their previous studies to determine the location of traffic counters in a network along with reducing the uncertainty in the estimation of O-D matrix and modification of the effects on data collection. While budget is limited, utilizing methods to improve the accuracy of O-D matrix is now frequently observed in the literature where entropy maximization method can even be implemented 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 way, license plate counts between the two cameras were considered as input, and then models are developed based on the least squared error and converted into a linear model that is seen as part of the O-D matrix. They approved that there is an estimation method which examines the daily fluctuations of O-D patterns in the highway network. This technique was later improved by portable license plate readers [17]. More improvement is also observed by considering probability while dynamic movement was studied in the region of Riga by Savrasovs and Pticina [18], to estimate the probability of detected destinations by recording vehicle type and plate number after dividing traffic into proper routes.

Network traffic density and uncertainty in traffic flow are also 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 are recommended to be applied in a dense network. They developed a nonlinear optimization model that includes a dynamic dispersion parameter followed by calculating the minimal squares of total matrix errors and repetition of service matching models to achieve 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 are 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 resulted by selected those arcs which had the basic role in arc-path matrix-vector. Statistical methods are also utilized for estimating the O-D matrix through the numerical data. Static and dynamic estimation of the O-D matrix, estimated matrix reliability, determining a set of traffic stations, the number of links need 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 on the installation cameras and traffic control sensors make significant improvements for determining the optimum arcs for flow counting and eventually estimating 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 the 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 main problem which is resource constraint, makes the above-mentioned methods not always feasible. In this way, it is essential to achieve the optimal number of traffic counter stations and their position in the network [21].

Using mobile phones is increasingly used for gathering data on traveling cars over the network to estimate the O-D matrix. Shahadat 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 each mobile tower in time window intervals. They revealed that using mobile phone data is more economic than traditional survey methods developed based on traffic counters. Combination of route choice and traffic counter data is also utilized to estimate O-D matrix in congested networks [23] where the basic principles of stochastic user equilibrium are assumed for route selection. The other studies in this field have been conducted to gain more accurate estimation of trip purposes such as Lauren et al. [24] to estimate the purposes of trips including home, work and the 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 the researches on estimating O-D matrix using traffic counters data include the operational status and practical techniques such as mathematical modeling, simulation, statistical methods but license plate recognition cameras and other ITS devices have been used in the recent years. Methods are usually based on traffic volume of all vehicles detected by their plates over the network, but the fundamental issues are extracted from two different perspectives comparing to the previous studies. The first is to consider the concept of vehicle tracking using license plate detection cameras across the inter-city network and passing through the selected paths. The second is to estimate the O-D matrix for freight transportation by just one or two detection points, followed in the present research work even trucks maybe detected by more cameras. So, a two-stage mathematical model is presented. The routes for passing of vehicle fleets on the wide network are determined at the first stage followed by license plate detection equipment used for guessing the origin and destinations. The second part is designed as an iterative procedure which moves gradually the first solution to the best by minimizing the deviation between observed and estimated traffic volumes on the network links. In other words, the main purpose of this article is to estimate the O-D demand in freight transport network, using 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 details in the second section. To explain how the procedure works, an illustrative example is then discussed in the third section followed by more discussions 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 origin-destination matrix. So, this section is divided into two parts. At the first part, an 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 to assign non-zero elements to O-D matrix known as basic variables in mathematical programming[26] where the procedure is proposed to gradually update them.

2.1. Iterative procedure

The iterative procedure includes some stages which are discussed as follows. At the initial stage, the network specifications including nodes, links or edges which connect network nodes, and distances between nodes are set. The O-D pairs known as candidates are then defined according to transport stats. Each O-D pairs should include the number of vehicles passing from origin to destination. All 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) obtains all links’ involvements for all O-D pairs which are used for solving the mathematical model in the next stage.

Input data is defined as a unique pattern proposed specifically in this research work. Defining this pattern is this research's novelty 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 the other links. The first and the last detection times are investigated, considering the time interval and assigned to each truck. It defines the number of vehicles which are observed at the first time on the first link and at the last time on the last link. So, the structure of input data is a four-dimension table including the first and the last passing links, each is separately defined by two start and end nodes. In this case, objective function is to minimize the difference between the observed and estimated number 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 is performed by updating variables by upper and lower limits for estimated O-D matrix elements until no significant changes observed in two continuous iterations. The root of average square errors of assigned and observed transport demand 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 an overall view of the proposed procedure where iterative steps are depicted by thick arrows.

{Please insert figure 1 here.}

2.2. Developing Mathematical Model

- **Basic Concept**
The basic concept in developing the mathematical model for estimating the O-D matrix lies in tracking and detecting the movement of trucks on the network arcs. As shown in figure (2), it is assumed that the vehicles that are moving from origin node “o” to 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 vehicle direction movement is also depicted by an arrow over equipped links. It is assumed that the location of equipped links is already known means locating arcs for installing the license plate system is out of this research work.

{Please insert figure 2 here.}

- **Notations and Parameters**
Parameters “i” and “j” are nodes 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 node and end node, respectively. In the intercity road networks, roads are commonly used in both directions, so the return arc is defined by similar properties defined as equation (1).

\[
(i, j) \quad \text{and} \quad (j, i) \in G
\]  

\(L_{ij}\) 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 equation (2).

\[
L_{ij} = L_{ji} \in G
\]

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 ones. Following the basic concept, the number of vehicle fleets are known to be detected through equipped arcs is defined as the following parameters:

\[
RP_{ij1}^{i2j2} : \text{The number of vehicles moving from their origins to destinations detected at the first time on arc } (i_1-j_1) \text{ and at the last time on arc } (i_2-j_2).
\]

Because there are many short distance 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 models’ performance in short distance O-D pairs, it is possible to consider vehicles detected by just one equipped link as defined by equation (3).

\[(i_{1,j_{1}})=(i_{2,j_{2}}) \text{ for cases in which vehicles passed one equipped arc} \quad (3)\]

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

\[M_{od}: \text{The number of vehicles currently move from origin “o” to destination “d”}. \]

\[\text{Given the above parameters are known, two decision variables are defined. The first variable is to determine the shortest routes over the network and the second variable is to determine the trucks pass through O-D pairs. Since the first part of the model is to determine all routes for all O-D candidates; the first decision variable is defined by equation (4) as a binary variable.} \]

\[X_{ij}^{od} = 1 \text{ if link (i,j) is located on the shortest path from origin O to destination D,} \]

\[\text{Otherwise is 0.} \quad (4)\]

The second decision variable which determines the assigned O-D matrix elements is defined as below.

\[Y_{od}: \text{The number of vehicles which move from origin “o” to destination “d”}. \]

- **Modeling procedure**

It is assumed that traffic volume does not have effects on vehicles’ 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; Castillo et al [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 in the initial stage can be defined by equation (5), where \(Z_1\) is the sum of the total traveled distance for all origins to destinations. This equation guarantees that all drivers select their shortest routes from their origins to corresponding destinations.

\[\text{Min } Z_1 = \sum_{(o,d)\in CN} \sum_{(i,j)\in G} L_{ij} \times X_{ij}^{od} \quad (5)\]

\[\text{There is one constraint in the initial stage which is to keep the seamless property of the route from origin to destination. This constraint is formulated by equation (6) where } E_x(j) \text{ is the set of exiting arcs from and } E_n(j) \text{ is the set of entering arcs to node “j”. More details in general modeling are available at [26] and in practical usage at[28].}\]

\[\sum_{i\in E_x(j)} X_{ij}^{od} - \sum_{i\in E_n(j)} X_{ij}^{od} = \begin{cases} 1 & \text{if } j = o \\ -1 & \text{if } j = d \\ 0 & \text{if } OW \end{cases} \forall j \in G \text{ and } (o,d) \in CN \quad (6)\]
The first part of the proposed procedure determines the links which are located in the specific O-D pairs. Therefore, the result shows that if the link \((i,j)\) is located on the path \((o,d)\) or not. From now on, the symbol \(X_{ij}^{od}\) is not considered as a variable and it is used as a parameter in the second model 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 transmitted from origin “o” to 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 Equation (7). The remarkable point is that the equation (7) covers just equipped arcs, so involved or relevant arcs are represented by equation (8). In fact, equation (8) guarantees avoiding duplications when trucks are detected by one camera installed over the link \((i1,j1)\) which is also recounted by equation (3). In this case, when a link is counted twice, the equation (8) deletes the redundant truck counting.

\[
Min \quad Z_2 = \sum_{(i,j)\in G} \left| \sum_{(o,d)\in CN} X_{ij}^{od} \times Y_{od} - FM_{ij} \right| 
\]

\[(i,j) = (i_1,j_1) \cup (i_2,j_2) \in G \quad (7)\]

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

\[
Min \quad Z_2 = \sum_{(i,j)\in G} PS_{ij} + NG_{ij} 
\]

\[
\sum_{(o,d)\in CN} X_{ij}^{od} \times Y_{od} - FM_{ij} - PS_{ij} + NG_{ij} = 0 \quad \forall (i,j) \in G \quad (9)\]

\[
\sum_{(o,d)\in CN} X_{ij}^{od} \times Y_{od} - FM_{ij} - PS_{ij} + NG_{ij} = 0 \quad \forall (i,j) \in G \quad (10)\]

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 equations (11) and (12). Equation (11) satisfies the number of vehicles which pass through the first equipped arc and equation (12) satisfies the number of vehicle candidate pairs for each O-D candidate pairs.

\[
\sum_{(o,d)\in CN} X_{i_1j_1}^{od} \times Y_{i_1j_1} = \sum_{(i_1,j_1)\in G} R_{i_1j_1}^{i_2j_2} \quad \forall (i_1,j_1) \in G \quad (o,d) \in CN \quad (11)\]

\[
\sum_{(o,d)\in CN} X_{i_2j_2}^{od} \times Y_{i_2j_2} = \sum_{(i_2,j_2)\in G} R_{i_2j_2}^{i_1j_1} \quad \forall (i_2,j_2) \in G \quad (o,d) \in CN \quad (12)\]

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 the number of equipped arcs multiply to 3 according to equations (10), (11) and (12). In this case, the number of non-zero variables, named also as basic variables, is less than the number of constraints. Therefore, the number of O-D pairs is bigger, so 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 moves traveled assignments to the specific values, step by step. The specific value is considered as old O-D pairs value defined as parameter $F_{ij}$ at the previous sub-section. In the iterative procedure, the obtained results are compared to the previous iteration results. By comparing the results, two cases may occur as follows;

**Case one:** After running the mathematical model, traffic volume assigned to the specific link is less than the old O-D or zero. In this case, a lower bound is necessary to restrict assigned traffic volume for the link. Figure (3) depicts the first situation in which assigned traffic volume is less than the O-D assigned value used for validation. The assigned value may be the old O-D pairs. The horizontal axes show the different values of O-D investigated in the solving procedure. In addition, an upper bound is also required to restrict the assigned traffic volume since the mathematical model may obtain unusual traffic volume. Following the above discussion, the movement toward better solution is satisfied by equation (13) where $\alpha$ is a coefficient for movement speed, $\beta$ is an acceptable coefficient to increase the lower bound, $Y_{od}^n$ is traffic volume assigned to each O-D pairs at iteration “n”, $RL_{od}^n$ is the amount of lower bound, and eventually $\gamma$ is an acceptable coefficient to raise the upper bound of assigned traffic volume to each O-D pairs. The next iteration is done to update all the above variables by equations (14) and (15).

\[
\text{If } Y_{od}^n \leq M_{od}; \\
\begin{align*}
RL_{od}^n &= \text{Min} \left[ Y_{od}^n + \alpha (M_{od} - Y_{od}^n), \quad \beta M_{od} \right] \\
Y_{od}^{n+1} &\geq RL_{od}^n \quad \forall od \in OD \quad \text{and} \quad Y_{od}^n \leq M_{od} \\
Y_{od}^{n+1} &\leq \gamma M_{od} \quad \forall od \in OD \quad \text{and} \quad Y_{od}^n \leq M_{od}
\end{align*}
\]

(13)

(14)

(15)

**Case two:** After running the mathematical model, the assigned traffic volume is much greater than the old O-D. In this case, an upper bound is necessary to limit assigned traffic volume for the specific link. Figure (4) depicts the first state in which assigned traffic volume is greater than the desired situation. In addition, a lower bound is also necessary to restrict the assigned traffic volume since the mathematical model may obtain unusual traffic volume. The movement toward better solution is satisfied by equation (16) where all coefficients and variables are the same as case one and $RU_{od}^n$ is the amount of upper bound of assigned traffic volume of O-D pairs. The next iteration is done to update all the above variables by equations (17) and (18).
If $Y_{od}^n \geq M_{od}$:

$$RU_{od}^n = \text{Max} \left[ Y_{od}^n - \alpha(Y_{od}^n - M_{od}), \gamma M_{od} \right]$$

(16)

$$Y_{od}^{n+1} \leq RU_{od}^n \quad \forall od \in OD \quad \text{and} \quad Y_{od}^n \geq M_{od}$$

(17)

$$Y_{od}^{n+1} \geq \beta M_{od} \quad \forall od \in OD \quad \text{and} \quad Y_{od}^n \geq M_{od}$$

(18)

3. Illustrative example

In order to understand how the proposed model works, an illustrative example is discussed. It is 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 as well as the connections are depicted as single lines mean 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 tabulated in Table (1) contains the length of the shortest path for each O-D pairs. For example, the first row shows data for O-D pairs from node 1 to node 5. The current travel O-D demand is 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 it is shown in the last column. Other rows show more information for the remaining O-D pairs. It is also assumed that arcs (1,3), (2,4), (2,7), (3,2), (3,5), (2, 4), (6,5) and (7,5) are equipped by license plate reader systems. The vehicles which move from origin “o” to destination “d” were firstly detected in the arc $(i_1, j_1)$, and for the last time in the arc $(i_2, j_2)$. All traveled data on vehicles passing arcs are tabulated in Table (2). For example, the first row indicates that there were 500 vehicles which move from 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 more for the other links.

Please insert table 1 here.
To estimate the O-D matrix elements, the proposed procedure and mathematical model are utilized using the well-known software of GAMS where the moving coefficients “α” is set to be 0.05, β is set to be 0.75 and γ is set to be 1.25. The results are represented in Table (3). As shown in 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 but some are zero and some are greater than the old demands. For example, in the third row, the current demand for O-D pairs 3 to 4 is 200 but it is assigned by zero at the initial stage. The root mean absolute error is also calculated as 211. As procedure evolves, the assigned O-D pairs move gradually to a steady state where at the third iteration, assigned O-D pairs for (3 to 4) changed to 108. Running the proposed procedure improves criterion which means that all assigned O-D pairs reached to steady state in the 10th iteration. No more iteration is required because no change is made after iteration 10. So, the final O-D pairs are now estimated in iteration 10. In order to check the convergence rate of the procedure, Mean Absolute Error (MAE) is used to calculate the difference between the observed and the 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 in 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 area of these provinces is 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 adjust for running the model. To improve data quality, border points are defined as origin or destination nodes. For example, total transport demand from origin or to destination for the provinces of Khorasan Razavi, North Khorasan, South Khorasan and Semnan to western provinces of Iran, is incorporated to Garmsar, 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 entered to or departed from the area under the study. More explanations on how to define border nodes are available at [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 is necessary for solving this problem. 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.

(Please insert table 4 here.)

Utilizing the proposed procedure obtains the estimated O-D Matrix over the network as shown in Table (5) for different iterations. In addition to the observed and assigned O-D values, the root mean square error is also calculated. To compare the model results and the observed trucks assigned to the links, the well-known criterion of Root Mean Square Error (RMSE) is investigated. The difference between the observed and the assigned trucks is calculated. This criterion shows how the model converges to the real situation. As shown in 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.

(Please insert table 5 here.)

Looking at the root mean square errors reveals that the convergence speed for the initial steps is higher than those for the final steps as shown in figure 7. It means that the procedure can be used in a few steps if the network is large-size and accuracy is not strongly important.

(Please insert figure 7 here.)

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 parameters performances. To check the performance, different ways are examined. For this purpose, the upper and lower limits are considered for the range of response as values of β and γ for different values of 0.75, 0.85 and 0.95, respectively, with the corresponding values limit above 1.25, 1.15 and 1.05. The above coefficients show that the upper and lower bounds have the same interval where they are adjusted from 1. However, it is not necessary. So, researchers may use different intervals. Moving speed is also checked by values of 0.01, 0.02
and 0.05. Using the coefficient $\beta=0.75$ means 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 $\alpha$ which justifies the speed of converging to the upper or lower bounds. Results summarized in table (6) reveal the following conclusion remarks:
- The error of the proposed model is decreased adding more iterations.
- O-D pairs reach to steady states after 200 iterations. This indicates 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 compared with upper ones.
- Upper and lower limits make the results to be closer to the existing situation.

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

(Please insert table 6 here.)

More discussions are also made on the convergence speed obtained by the proposed procedure. In order to check the above consideration, the obtained root mean square errors are depicted in figure 8 for $\beta=0.75$ and $\gamma=1.25$, using iteration numbers. As shown, the procedure finally converges to the same solution but the coefficient $\alpha$ plays an important role throughout the calculation process.

(Please insert figure 8 here.)

5. Summary and Conclusion
The O-D matrix estimation is one of the most important steps in transport planning. Since electronic devices which 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 are 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 is to minimize the residuals obtained from the observed and assigned number 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.
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 gathering data of current O-D matrix, for one-year duration, results were examined for 145 links which were equipped before for the 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, is capable to accurately estimate the elements of the O-D matrix. Since the installation of electronic devices is being widespread, further researches are recommended to focus on considering traffic volumes over the links or investigating the local or temporary road closing programs mainly imposed by transport authorities over the network. In these cases, the connectivity reliability of link can be considered to routing problem stage which is under authors’ studies for the future.

References


Network specifications including nodes, arcs, and distances

Developing mathematical model

Identify OD pair Candidates

Obtaining shortest paths for all OD candidates

Determining the number of vehicles crossing equipped arches

Determining the movement step and upper and lower limits parameters

Solving mathematical model for estimating OD matrix

Obtaining shortest paths for all OD candidates

Equipped link

Equipped link

Moving Direction of Vehicles

Updating the upper and lower limits of O-D candidates

Yes

No

Stopping criterion met?

Explore final O-D travel demands

Figure 1. Overall view of the proposed procedure

Figure 2. The general pattern of moving vehicles detected over the equipped links
Figure 3. Movement direction when the assigned O-D is less than the old value

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

Figure 5. Illustrative network structure

Figure 6. An overall view of the case study highlighted in yellow
Figure 7. Root Mean Square Errors Based on the Number of Iterations

Figure 8. Root Mean Square Errors Based on the Convergence Coefficient of $\alpha$
Table 1. Data assumed for illustrative example

<table>
<thead>
<tr>
<th>Row</th>
<th>Origin</th>
<th>Destination</th>
<th>O-D demand</th>
<th>Shortest path</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>5</td>
<td>500</td>
<td>1-&gt;3-&gt;5</td>
<td>135</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>6</td>
<td>650</td>
<td>1-&gt;2-&gt;4-&gt;6</td>
<td>185</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
<td>200</td>
<td>3-&gt;2-&gt;4</td>
<td>115</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>7</td>
<td>350</td>
<td>3-&gt;2-&gt;7</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>2</td>
<td>300</td>
<td>5-&gt;3-&gt;2</td>
<td>135</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>2</td>
<td>450</td>
<td>6-&gt;4-&gt;2</td>
<td>145</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>5</td>
<td>400</td>
<td>6-&gt;5</td>
<td>80</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>7</td>
<td>450</td>
<td>6-&gt;5-&gt;7</td>
<td>140</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>6</td>
<td>250</td>
<td>7-&gt;5-&gt;6</td>
<td>140</td>
</tr>
</tbody>
</table>

Table 2. Equipped links involved on passing vehicles (Illustrative example)

<table>
<thead>
<tr>
<th>Row</th>
<th>The first arc that records the plaque</th>
<th>The last arc that registers the plaque</th>
<th>Number of vehicles</th>
<th>Row</th>
<th>The first arc that records the plaque</th>
<th>The last arc that registers the plaque</th>
<th>Number of vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(1,3)</td>
<td>(3,5)</td>
<td>500</td>
<td>5</td>
<td>(3,2)</td>
<td>(3,2)</td>
<td>300</td>
</tr>
<tr>
<td>2</td>
<td>(2,4)</td>
<td>(2,4)</td>
<td>850</td>
<td>6</td>
<td>(4,2)</td>
<td>(4,2)</td>
<td>450</td>
</tr>
<tr>
<td>3</td>
<td>(3,2)</td>
<td>(2,4)</td>
<td>200</td>
<td>7</td>
<td>(6,5)</td>
<td>(6,5)</td>
<td>850</td>
</tr>
<tr>
<td>4</td>
<td>(3,2)</td>
<td>(2,7)</td>
<td>350</td>
<td>8</td>
<td>(7,5)</td>
<td>(7,5)</td>
<td>250</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Row</th>
<th>Origin</th>
<th>Destination</th>
<th>Existing Demand</th>
<th>Iteration 0</th>
<th>Iteration 3</th>
<th>Iteration 5</th>
<th>Iteration 10</th>
<th>Iteration 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>5</td>
<td>500</td>
<td>200</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>6</td>
<td>650</td>
<td>1050</td>
<td>942</td>
<td>886</td>
<td>850</td>
<td>850</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
<td>200</td>
<td>0</td>
<td>108</td>
<td>164</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>7</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>350</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>2</td>
<td>300</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>2</td>
<td>450</td>
<td>450</td>
<td>450</td>
<td>450</td>
<td>450</td>
<td>450</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>5</td>
<td>400</td>
<td>0</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>7</td>
<td>450</td>
<td>850</td>
<td>450</td>
<td>450</td>
<td>450</td>
<td>450</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>6</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
</tr>
<tr>
<td>MAE*</td>
<td></td>
<td></td>
<td></td>
<td>211</td>
<td>53</td>
<td>34</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

* Mean Absolute Error
### Table 4. The number of vehicles passed over the first and the last links

<table>
<thead>
<tr>
<th>Row</th>
<th>First Detecting Link</th>
<th>Last Detecting Link</th>
<th>Number of vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(1179, 1178)</td>
<td>(1307, 1001)</td>
<td>13965453</td>
</tr>
<tr>
<td>2</td>
<td>(1196, 1194)</td>
<td>(1141, 1121)</td>
<td>4842134</td>
</tr>
<tr>
<td>3</td>
<td>(1196, 1194)</td>
<td>(1235, 1131)</td>
<td>2417879</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>2509</td>
<td>(1080, 1089)</td>
<td>(1025, 1053)</td>
<td>51</td>
</tr>
<tr>
<td>2510</td>
<td>(1035, 1171)</td>
<td>(1177, 1178)</td>
<td>51</td>
</tr>
<tr>
<td>2511</td>
<td>(1022, 1047)</td>
<td>(1081, 1079)</td>
<td>51</td>
</tr>
</tbody>
</table>

### Table 5. Estimated O-D Matrix in different iterations

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

### Table 6. Sensitivity analysis results using different parameters

<table>
<thead>
<tr>
<th>Row</th>
<th>β</th>
<th>γ</th>
<th>α</th>
<th>Iteration 0</th>
<th>Iteration 10</th>
<th>Iteration 20</th>
<th>Iteration 50</th>
<th>Iteration 75</th>
<th>Iteration 100</th>
<th>Iteration 200</th>
<th>Iteration 500</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.75</td>
<td>1.25</td>
<td>0.01</td>
<td>624607</td>
<td>512321</td>
<td>451183</td>
<td>329621</td>
<td>198005</td>
<td>80884</td>
<td>25645</td>
<td>25435</td>
</tr>
<tr>
<td>2</td>
<td>0.02</td>
<td></td>
<td>0.01</td>
<td>624607</td>
<td>450695</td>
<td>365945</td>
<td>197302</td>
<td>80105</td>
<td>25732</td>
<td>25435</td>
<td>25435</td>
</tr>
<tr>
<td>3</td>
<td>0.05</td>
<td></td>
<td>0.01</td>
<td>624607</td>
<td>333578</td>
<td>196179</td>
<td>50417</td>
<td>25653</td>
<td>25835</td>
<td>25835</td>
<td>25835</td>
</tr>
<tr>
<td>4</td>
<td>0.85</td>
<td>1.15</td>
<td>0.01</td>
<td>624607</td>
<td>504263</td>
<td>452086</td>
<td>331343</td>
<td>198810</td>
<td>72552</td>
<td>17399</td>
<td>17399</td>
</tr>
<tr>
<td>5</td>
<td>0.02</td>
<td></td>
<td>0.01</td>
<td>624607</td>
<td>456505</td>
<td>370350</td>
<td>198407</td>
<td>72886</td>
<td>20160</td>
<td>17669</td>
<td>17669</td>
</tr>
<tr>
<td>6</td>
<td>0.05</td>
<td></td>
<td>0.01</td>
<td>624607</td>
<td>336288</td>
<td>198916</td>
<td>46356</td>
<td>17386</td>
<td>17669</td>
<td>17669</td>
<td>17669</td>
</tr>
<tr>
<td>7</td>
<td>0.95</td>
<td>1.05</td>
<td>0.01</td>
<td>624607</td>
<td>515161</td>
<td>453314</td>
<td>331310</td>
<td>197597</td>
<td>67676</td>
<td>7735</td>
<td>7735</td>
</tr>
<tr>
<td>8</td>
<td>0.02</td>
<td></td>
<td>0.01</td>
<td>624607</td>
<td>455118</td>
<td>369860</td>
<td>197622</td>
<td>68434</td>
<td>10166</td>
<td>7678</td>
<td>7678</td>
</tr>
<tr>
<td>9</td>
<td>0.05</td>
<td></td>
<td>0.01</td>
<td>624607</td>
<td>334878</td>
<td>197899</td>
<td>38061</td>
<td>7694</td>
<td>7678</td>
<td>7678</td>
<td>7678</td>
</tr>
</tbody>
</table>
Biographies

Dr. Mohammad Teimouri received his undergraduate degree from Yazd University in 1997. He then received his Master degree in Socio-economic Systems Engineering from Tarbiat Modares University Tehran, Iran in 2000. He holds PhD in Industrial Engineering. Dr. Mohammad Teimouri received his Ph.D from Department of Industrial Engineering, Branch of Tehran South, Islamic Azad University. His main research area is Intelligent Transportation Systems. Currently, he is vice president of Iran’s Road Maintenance and Transportation Organization.

Dr. Majid Sheikhmohammady received his Ph.D in Systems Design Engineering from university of Waterloo, Canada, in January 2009. He was also a Post Doctoral Fellow in the department of Systems Design Engineering, University of Waterloo, in 2009. He is currently a faculty member in 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 analyse 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 defence science.

Dr. 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 in the Department of Industrial and Systems Engineering, Tarbiat Modares University and has been active in 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. As solution methodologies for real world engineering design problems, he has introduced several intelligent optimization procedures, which inspire 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”.

Dr. Amir Abbas Shojaei holds PhD in Industrial Engineering. He received his Ph.D from Department of Sciences and Researches, Islamic Azad University Tehran, Iran n 2003. Currently, he is a faculty member in the Department of Industrial Engineering, Branch of Tehran South, Islamic Azad University. His main research area includes transportation planning, production policies, and control project.