OPTIMAL SENSOR LOCATION AND ORIGIN–DESTINATION MATRIX OBSERVATION WITH AND WITHOUT SENSORS ON UNCONGESTED NETWORKS

Hadi KARIMI[1*], Seyed-Nader SHETAB-BOUSHEHRI[2], Ali ZEINAL HAMADANI[3]

[1, 3]*Dept of Industrial and Systems Engineering, Isfahan University of Technology, Iran*
[2]*Dept of Transportation Engineering, Isfahan University of Technology, Iran*

Abstract. The Origin–Destination (O–D) matrix, is an important information in transportation planning and traffic control. Rapid changes in land use, particularly in developing countries, have been and are on an increase, which makes the estimation and observation of this matrix more significant. The objective of this paper is to observe O–D matrix under two scenarios. In the first scenario, it is assumed that the traffic network is equipped with path-ID sensors. In this situation, the goal is to determine the optimal number and location of these sensors in the network, where by applying collected information through these sensors, the O–D matrix is observed. Because path-ID sensors are not available in many cities, in the second scenario the interview alternative is proposed in order to observe O–D matrix. The interview method has encountered some restrictions. Several mathematical programming models have been developed to overcome these restrictions. To illustrate these proposed methodologies, they are applied in the Nguyen–Dupuis transportation network and the results are analysed. By applying the model on the intercity road network in the Province of Isfahan (Iran), a large network, the efficiency of these proposed models is demonstrated. Finally, some conclusions and final recommendations are included.

Keywords: origin–destination matrix, observability problem, network sensor location problem, uncongested networks, path-ID sensors, Province of Isfahan.

Introduction

Identifying accurate Origin–Destination (O–D) matrix is one of the most important and challenging problems in transportation planning and traffic management (Kim *et al.* 2018). The elements of O–D matrix represent how trips are distributed among different traffic zones in the area under study and reflect traffic flow information thereof in a particular time period (Pravinvongvuth 2007). This matrix is applied as a critical source of input information in transportation models in order to draw and assess various decisions in transportation planning within short-term managerial issues to long-term strategic objectives' range (Chootinan *et al.* 2005). Given its significance, O–D matrix has become a major issue of study with the objective to propose different estimation or observation methods, based on the type and quality of traffic flow information collected from the network, prior information and passenger route choice behaviour. A large portion of information required in estimating or observing the ma-

trix is obtained through survey methods like interviewing network users, plate scanning, the vehicle count, etc. However, this is a costly procedure to construct O–D matrix, which must also be updated in short intervals. Thus, beginning the past three decades, many attempts have been and are being made to improve the available methods. Recent transportation studies as to traffic flow information collection from networks are subject to two problem sets: (1) problems where collecting network flow information does not yield a unique solution for the O–D matrix, known as the estimation problem in extant literature. An overview of O–D estimation problem and a brief survey of recent solutions are presented in De Grange *et al.* (2017) and (2) problems where flow information gathered from the network leads to a unique solution for the O–D matrix, known as the observability problem. The focus here is on observability problem.

*Corresponding author. E-mail: *h.karimidehnavi@in.iut.ac.ir*

Observations of O–D flow data are rare. In recent years, there exist studies on how to explore and extract such data from traffic sensors. Four types of traffic sensors based on their functionality are categorized by Gentili and Mirchandani (2012): (1) counting sensors capable of counting vehicle count passing through one or multiple lines, which would lead to monitor lane performance like: speed, density, occupancy and flow rates, (2) image sensors, as a fixed camera mounted on a tall building or a pole measuring flows at an intersection or a link. The moving vehicles in the scene are recognized after processing the images, (3) vehicle-ID sensors, through which a vehicle-ID and the time the vehicle activates these sensors can be identified on the network and (4) path-ID sensors placed on network links to measure the flow of passing vehicles. This type of sensor on a given arc of the network measures the flow of the paths (routes) to which the arc belongs. Although path-ID sensors are currently exclusive to certain vehicles like buses and trucks, by devising the necessary infrastructure, they can be applied in recording paths of all vehicles, thus making the observability problem easy to solve. Unlike other sensors, not enough studies are run on path-ID sensors.

Based to the above mentioned, the first objective of this paper is to observe the flow of all O–D pairs by applying path-ID sensors. In this context, a model is proposed which would determine the optimal number and location of these sensors in the network, where by applying collected information through these sensors, the O–D matrix is observed (determined uniquely).

In practice, the data obtained from counting sensors often are not capable of finding a unique solution for O–D flows, because the number of known variables (link flows) exceeds the number of unknown variables (O–D flows). These data are frequently applied in estimating or updating the O–D matrix. By applying image sensors the turning ratio in an intersection can be determined. Although vehicle-ID and path-ID sensors are capable to record path information, due to technical and economic restrictions (especially in developing countries), the prerequisite required in applying these sensors is not available in identifying the path of all O–D pairs, thus a major challenge issue in current era.

Due to absence of these sensors, the second objective of this paper is to observe the flow of all O–D pairs by data, which is collected by interviewers from the users of the paths during their trips. Put another way, by stopping vehicles at selected links in the network and inquiring about their origins and destinations, the required information can be collected. As to this approach, there exist two challenges: (1) due to the presence of multiple paths for each O–D, it is extremely complicated to manually identify and record user paths in large real networks and (2) given the volume of traffic passing through network links, it is impossible to stop and interview every vehicle, therefore, interviewers are forced to select a limited number (i.e. a sample) of network users, generalization of which the sample results leads to errors in O–D matrix. In this paper, many mathematical programming models are developed to overcome these challenges.

The remainder of this paper is organized as follows: Section 1 briefly surveys the literature dealing with the principal aspects of our model; Section 2 presents the methods for solving the observability problem of O–D matrix using path flow information. Section 3 reports the results obtained from running the proposed models on both the Nguyen–Dupuis network and the real road network in the Province of Isfahan (Iran). Finally, conclusions and further studies are given in conclusion Section.

1. Literature review

The O–D matrix, is an essential source of information applied as input in transportation planning and traffic management, the observation of which has gained major interest among researchers in the recent years. In general, network flow observability problem is involved with applying information from observed flows in the network in order to determine the subset of flows, which can be calculated (Rinaldi, Viti 2017). A set of flows (e.g. link, path and O–D), would provide the possibility for the unobserved flows to be observed (Castillo *et al.* 2012).

Solutions to the observation problem are presented in the available literature by exploiting the fundamental equation among the three sets of flow variables (link, route and O–D), yield from conservation laws (Rinaldi, Viti 2017). Equation (1) known as conservation laws, convey the most relevant information in network flows:

$$v_a = \sum_{r \in R} \beta_{ar} \cdot f_r, \ \forall a \in A;$$

$$t_w = \sum_{r \in R} \rho_r^w \cdot f_r, \ \forall w \in W, \tag{1}$$

where: β_{ar} is an element of the link-route incidence matrix, equal to one if path r in the set of network paths R passes through link a, otherwise equal to zero; ρ_r^w is an element of the route O–D incidence matrix, where if path r belongs to O–D pair w then ρ_r^w is equal to one, otherwise is zero; v_a, f, t_w are the volume of link a, flow of path r and the O–D flow w, respectively.

The first set in Equation (1) indicates that the volume passing through link a equals the sum of path flows passing the link, while the second set in Equation (1) indicates that the flow of O–D pair w (i.e. t_w) equals the sum of path flows in O–D pair.

If unobserved flows in the network are determined through the information from a set of observed flows and the equation above, then full observability is established (Xu *et al.* 2016).

The observability problems, depending on the set of available flows and the set of those which must be determined are categorized in three by Castillo *et al.* (2012): (1) observing link flows to determine path and O–D flows, (2) observing path flows to determine O–D and link flows and (3) observing a combination of O–D, path, and link flows to determine the flow of their unobserved sets.

Here, the first two categories are discussed in more detail. According to Castillo *et al.* (2012), observing link flows to determine path and O–D flows is often incapable of finding a unique solution for path flows (or O–D flows), because the number of known variables (link flows) exceeds that of the unknown variables (path flows). In this category, given the error in link flows, the above equation becomes incompatible, that is no set of path and O–D flows are met. On the contrary, by observing path flows, the O–D and link flows are easily obtained and the system remains compatible even in case of observing erroneous path flows. Thus, the only challenge to observe O–D matrix is to understand how the path flows are observed. In recent years, sensors are applied in recording traffic flows. Determining the type and location of sensors is the main issue in recording the users' path.

Many detailed trajectory datasets like Global Positioning System (GPS) traces from cell phones and vehicles, anonymized Call Detail Records (CDR) from cell phone providers, and data from arrays of Bluetooth and Wi-Fi detectors constitute the components of advanced technology. Consequently, several methods are developed for estimating O–D matrix (e.g. Li *et al.* 2018), for estimating travel time (e.g. Olia *et al.* 2017) and for providing traveller information (e.g. Caceres *et al.* 2008). Among these components, GPS data are applied more extensively due to their finer spatial and temporal resolution, which can be exploited to yield high quality route and travel time data by transportation researchers, planners, and modellers. In practice, due to the low penetration rates of these devices among transportation networks' users, excessive costs and the probability of statistical bias for specific social groups such as young people or business people (Bauer *et al.* 2018), most applications of GPS data are going through developing process as to modelling components, calibration and validation (Lee *et al.* 2016). The CDRs are the recent attempts made in improving travel surveys. In this process, the data are only recorded when the phone is active, and this function tends to be less accurate than GPS data, due to irregular with significant gaps in trajectory traces. Recently, some researchers are for applying path-ID sensors in order to gather traffic information from transportation networks.

The location of sensors within a network that meets specific objectives needs to be carefully selected, based on predefined objectives related to the effective collection of data and the subsequent estimation of traffic related information (Mitsakis *et al.* 2017). Observation of network flow is a subcategory of Network Sensor Location Problems (NSLP), the literature of which is related to O–D observability with a focus on algebraic and topologic features of network structure and connections where the objective is to install sensors to fully observe flow volumes on the network through linear independent equations with respect to network topology incidence (Viti *et al.* 2015). A polynomial time algorithm based on Gaussian elimination is devised by Hu *et al.* (2009) in order to determine the minimum number of counting sensors to be installed on links for observability of link flows, together with determining the necessary information from observed links to yield unobserved link flows. In this approach, a set of sensors are installed in a manner that the gained information can identify all unobserved flows uniquely. Elimination of path enumeration through algorithms is proposed by Ng (2012). In theoretical sense, these algorithms can take an exponential amount of Central Processing Unit (CPU) time in order to obtain an optimal solution. The number of sensors and their locations in upper bound as to completing link flow observability through the linearly independent paths is refined by Castillo *et al.* (2014). A graphical approach through topological tree-shaped characteristic is proposed by He (2013). In addition to link counts, the NSLP is studied in the presence of plate scanning technique in order to observe O–D and path flows (e.g. Castillo *et al.* 2013; Mínguez *et al.* 2010).

Next to the type, number and location of sensors, Equation (1) depend on β_{ar} and ρ_r^w parameters determined from network user route choice behaviour. Ortúzar and Willumsen (2011) classify methods of assigning trips to the network when the users have a common perception of travel costs (single-user class) with and without considering the stochastic effects. The simplest approach is the all-or-nothing assignment where congestion effects are disregarded and all drivers consider the same attributes in choosing their routes (i.e. stochastic effects are ignored). This unity in choosing a route is best assessed to sparse networks consisting of relatively few routes with vastly different costs. In contrast, there exist situations where drivers have dissimilar perceptions of costs and use a combination of criteria such as distance, travel time, path aesthetics, and safety to minimize travel costs. In this context, the stochastic effects need to be considered in path selection. In these models, in addition to the shortest path, drivers may choose alternative routes within many. The most well-known method is Dial's traffic assignment model, where the flow corresponding to each path of a particular OD is obtained according to the following equation:

$$t_r^w = \frac{t^w \cdot \exp\left(-\theta \cdot C_r^w\right)}{\sum_r \exp\left(-\theta \cdot C_r^w\right)}, \tag{2}$$

where: t^w is the flow of O–D pair w; t_r^w is the flow of path r in O–D pair w; C_r^w is the travel cost of path r in O–D pair w and θ is the parameter controlling the distribution of trips among paths.

In models with capacity restraints, network users choose their routes by following equilibrium conditions and considering the fact that the cost of each trip at each link varies with the flow of the given link. In some studies it is revealed that there exist factors like physical properties of the infrastructure and the users' detailed personal preferences, which are important in choosing routes (Földes, Csiszár 2015).

2. Methods

According to literature review section, it is clear that unlike other sensors, path-ID sensors have received relatively little research attention. Thus, in this section by considering access to path-ID sensors, a model is developed for observing O–D matrix. On the other hand, due to technical and economic restrictions (especially in developing countries), applying sensors like path-ID, GPS, CDRs, etc., which are capable of determining users' paths is facing some challenges. Consequently, in this paper a model is proposed for observing O–D matrix without applying these sensors. In other words, in this study two main situations of access to sensors (path-ID) and no access to sensors are of concern.

2.1. Situation 1: access to path-ID sensors

Here, the required technology and infrastructure for path-ID sensors are assumed to be available in the uncongested transportation network – meaning that the network can be equipped with these sensors. Thus, for each vehicle detected by the sensors, the origin, destination, and path can be determined. It is also assumed that, network users in addition to lowest generalized cost to reach their destinations in choosing their paths, tend to consider other factors like aesthetics and safety. Bearing in mind the stochastic effects in users' choice of routes, traffic assignment becomes a multipath problem. Among which, Dial's approach is the most well-known in uncongested networks. By considering network user behaviour, Model 1 is proposed as:

Model 1

$$\min \sum_{a \in A} x_a \tag{3}$$

subject to:

$$\sum_{a \in A} \rho_{ra} \cdot x_a \geq 1, \ \forall r \in R; \tag{4}$$

$$x_a \in \{0,1\}, \ \rho_{ra} \in \{0,1\},$$

$$\forall a \in A, \ \forall r \in R. \tag{5}$$

This model is of a binary linear integer programming where, x_a is a binary variable associated with link a. This binary variable is one if the path-ID sensor is installed on link a, otherwise zero. The binary variable ρ_{ra} is one if path r is a reasonable path in the network passing link a, otherwise zero. R is the set of all reasonable paths in the network, and A and W are the sets of all links and all O–D pairs in the network, respectively.

Here, the objective function in Equation (3), determines the number and locations of path-ID sensors subject to Dial's assignment in an uncongested network. The set of constraints in Model 1, which assure that vehicle information is recorded by path-ID sensors on all reasonable paths of the network, when all paths are assured to be observed, observability conditions for O–D flows in the network are satisfied, are expressed in Equation (4).

The Branch and Bound (BB) and Gomory Cuts (GC) methods are adopted in problems with integer variables. BB is a common adopted method because of its efficiency in computation. A combination of BB and GC is introduced as an efficient technique named the Branch-Cut (BC), applied in this context (Castillo *et al.* 2001). The theory and methods for solving linear and integer programming is discussed in research by Schrijver (1998).

In practice, the necessary infrastructure for application of path-ID sensors in transportation networks is non-existing or limited to particular vehicles. In the following, information collection is proposed based on interviewers in the absence of path-ID sensors. Ortúzar and Willumsen (2011) have addressed the basic data collection and interview methods (e.g. household, roadside and cordon interview), where, it is revealed that understanding the use of data is one of the key steps in choosing the type of interview in transportation planning.

2.2. Situation 2: no access to path-ID sensors

Assuming that path-ID sensors are not available and passengers' paths are unknown, Situation 1 is converted to Situation 2. Here it is assumed that only origins and destinations are recoded by the interviewers. Accordingly, lack of knowledge on the paths selected by network users makes Model 1 incapable of distinguishing between path flows. This incapability is explained in the following example.

The transportation network in Figure 1 consists of five nodes and six links with two O–D pairs: $T_1 = (1, 4)$ and $T_2 = (1, 5)$. The traffic flow from Origin (1) may travel one of three paths $r_1 = \{a, b\}$, $r_2 = \{c, d, b\}$ and $r_3 = \{c, e\}$ to arrive in Destination (4). In addition, passing through $r_4 = \{c, f\}$, traffic flows from Origin (1) to Destination (5).

An optimal solution for the model identifies the links b and c as the locations for interview. It is assumed that users are asked only about their origins and destinations, where on link b, the sum of volumes moving through r_1 and r_2 from the T_1 are recorded while the sum of volumes travelling paths r_2 and r_3 from the same O–D is registered on link c. Obviously, by applying this information, T_1 O–D flow (i.e. sum of volumes in paths r_1, r_2 and r_3) cannot be uniquely calculated. Thus, allowing Model 2 to be proposed:

Model 2

$$\min \sum_{a \in A} x_a \tag{6}$$

subject to:

$$\sum_{a \in A} m_a^w \cdot \rho_{ra}^w \cdot x_a = 1, \ \forall r \in R^w, \ \forall w \in W; \tag{7}$$

$$x_a \in \{0,1\}, \ \rho_{ra}^w \in \{0,1\}, \ m_a^w \in R,$$

$$\forall a \in A, \ \forall w \in W, \ \forall r \in R^w. \tag{8}$$

Model 2 is a Mixed Integer Non-Linear Programming (MINLP) model, where, x_a is defined as what is in Model 1.
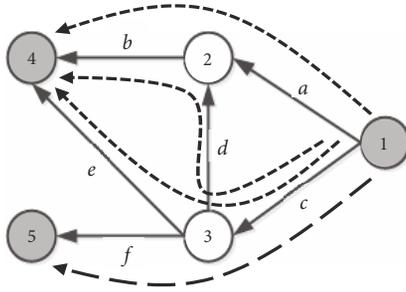
Figure 1. A hypothetical transportation network
illustrating overlap

The binary variable ρ_{ra}^w is one if path r of O–D pair w includes link a, otherwise zero; variable m_a^w is an unrestricted variable applied as a coefficient for constraint set (7) to assure the selection of interview locations yielding appropriate information applicable in determining O–D flows uniquely, in the network. In this model, objective function (6) minimizes the number of O–D interview locations, while constraint set (7) satisfies the path covering rule and guarantees that, by having a linear combination of the paths between any O–D pair, the pair's flow is obtained.

After Model 2 is solved for the hypothetical network in Figure 1, links a and c will be selected as the optimal solution for conducting interviews. It obvious that the volume on path r_1 from O–D pair T_1 is recorded on link a, while link c registers the sum of volumes on r_2 and r_3 from T_1 and r_4 from T_2. By applying this available information, T_1 and T_2 O–D flows as the sum of volumes on r_1 through r_3, and r_4 are determined uniquely, respectively. It is notable that $\{b, c\}$ is not a feasible solution in Model 2.

Because Model 2 is non-linear, solving real large-scale instances may be quite challenging. To overcome this drawback, Model 2 is rewritten as a Mixed Integer Program (MIP) in Model 3:

Model 3

$$\min \sum_{a \in A} x_a \qquad (9)$$

subject to:

$$\sum_{a \in A} \left(m_a^{+w} - m_a^{-w}\right) \cdot \rho_{ra}^w = 1, \ \forall r \in R^w, \ \forall w \in W; \qquad (10)$$

$$M \cdot x_a \geq m_a^{+w}, \ \forall a \in A, \ \forall w \in W;$$

$$M \cdot x_a \geq m_a^{-w}, \ \forall a \in A, \ \forall w \in W;$$

$$m_a^{+w} \geq 0, \ m_a^{-w} \geq 0, \ \forall a \in A, \ \forall w \in W; \qquad (11)$$

$$x_a \in \{0,1\}, \ \rho_{ra}^w \in \{0,1\},$$

$$\forall a \in A, \ \forall r \in R^w, \ \forall w \in W, \qquad (12)$$

where: m_a^w is an unrestricted variable, replaced with two non-negative variables of m_a^{+w} and m_a^{-w}; M is an arbitrarily large value. The remaining variables are defined as in Model 2. Constraint set (11) assure that m_a^{+w} and m_a^{-w} are selected in an unrestricted manner provided that link a is chosen to conduct interviews on.

As explained before, the large volume of traffic passing through network links would not allow all drivers to be interviewed, thus, only a fraction of the passengers are interviewed on a random basis. Generalizing the obtained results thereof leads to having some errors in the O–D matrix. Accordingly, the following procedure is applied to determine O–D flow variances.

2.3. Generalizing O–D information

For this purpose, let k represent the number of O–D pairs in the network where a proportion of the flows belonging to the 1-st…k-th O–D pairs pass through link a, and let λ be the fraction of network users whose origin and destination are enquired while travelling on link a. Assume that, firs, for each O–D pair, $n^1, n^2, …, n^k$ passengers are interviewed on link a and their O–D information is recorded, where $n^1 + n^2 + … + n^k = n$, second, the total number of passengers using link a is N, third, let $\tilde{X}^w$ is a random variable representing the number of instances observed from the flow of O–D w and p_w is the probability of interviewing a user from O–D w on link a. Consequently, the joint density function for observing different O–D flows on link a with a multinomial distribution is expressed as follows:

$$f\left(n^1, n^2, …, n^k; n, p_1, p_2, …, p_k\right) =$$

$$\Pr\left(\tilde{X}^1 = n^1, \tilde{X}^2 = n^2, …, \tilde{X}^w = n^w\right) =$$

$$\begin{cases} \dfrac{n!}{n^1!…n^k!} p_1^{n^1}…p_k^{n^k}, & \sum_{i=1}^{k} n^i = n; \\ 0, & \text{otherwise.} \end{cases} \qquad (13)$$

The percentage and estimated proportion of each O–D pair on link a is calculated as follows:

$$\lambda = \frac{n}{N}; \qquad (14)$$

$$\hat{p}_1 = \frac{n^1}{n};$$

$$\hat{p}_2 = \frac{n^2}{n};$$

$$…;$$

$$\hat{p}_k = \frac{n^k}{n}. \qquad (15)$$

In a sense that $\sum_i \hat{p}_i = 1$ holds true.

To construct the variance-covariance matrix resulting from this generalization, Fisher information function is applied. For k O–D flows passing through link a this function is expressed as follows:

$$I(\hat{p}) = \begin{bmatrix} \dfrac{n \cdot (\hat{p}_1 + \hat{p}_k)}{\hat{p}_1 \cdot \hat{p}_k} & \dfrac{n}{\hat{p}_k} & \cdots & \dfrac{n}{\hat{p}_k} \\ \dfrac{n}{\hat{p}_k} & \dfrac{n \cdot (\hat{p}_2 + \hat{p}_k)}{\hat{p}_1 \cdot \hat{p}_k} & \cdots & \dfrac{n}{\hat{p}_k} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{n}{\hat{p}_k} & \dfrac{n}{\hat{p}_k} & \cdots & \dfrac{n \cdot (\hat{p}_{k-1} + \hat{p}_k)}{\hat{p}_{k-1} \cdot \hat{p}_k} \end{bmatrix}. \qquad (16)$$

Given the properties of the Fisher information function, the inverse function $I(p)$ is the variance-covariance matrix, calculated as follows:

$$\Sigma = I^{-1}(\hat{p}) =$$

$$\frac{1}{n} \cdot \begin{bmatrix} \hat{p}_1 \cdot (1 - \hat{p}_1) & -\hat{p}_1 \cdot \hat{p}_2 & \cdots & -\hat{p}_1 \cdot \hat{p}_k \\ -\hat{p}_2 \cdot \hat{p}_1 & \hat{p}_2 \cdot (1 - \hat{p}_2) & \cdots & -\hat{p}_2 \cdot \hat{p}_k \\ \vdots & \vdots & \ddots & \vdots \\ -\hat{p}_k \cdot \hat{p}_1 & \hat{p}_k \cdot \hat{p}_2 & \cdots & \hat{p}_k \cdot (1 - \hat{p}_k) \end{bmatrix}, \quad (17)$$

where: the elements on the main diagonal are the variance in the proportion of O–D flows passing through link a.

By considering Equation (18) for the variance of a finite population,

$$\left(\frac{N-n}{N-1} \right) \cdot \frac{\sigma^2}{n} \qquad (18)$$

and by assuming $n = \lambda \cdot N$, $0 < \lambda \le 1$, the proportion of O–D flows Standard Deviation (SD) passing through link a is obtained through Equation (19):

$$\sqrt{\frac{1-\lambda}{\lambda} \cdot \frac{p_w \cdot (1 - p_w)}{N}}. \qquad (19)$$

Because the random variable X^w (volume of w-th O–D flow) on link a is predicted as $\hat{X}_a^w = N \cdot \hat{p}_w$ together with applying Equation (19), X_a^w SD on link a is obtained as follows:

$$SD(X_a^w) = \sqrt{\frac{1-\lambda}{\lambda} \cdot N \cdot \hat{p}_w \cdot (1 - \hat{p}_w)}, \qquad (20)$$

where: when all network users on selected links are interviewed (i.e. $\lambda = 1$) the SD is zero; however, as expected, this SD increases as λ approaches zero. Furthermore, $N \cdot \hat{p}_w$ (i.e. the counted flow in the w-th O–D on link a) is a constant value equal to n^w, therefor, the following is observed: as the proportion of w-th O–D flow increase on link a, the corresponding flow SD decreases.

One of the approaches in increasing λ parameter is to apply online O–D survey. Although online survey about O–D is run at a lower cost and time, the internet penetration rate, especially in developing countries, and lack of interviewers as process facilitators constitute the main challenges of this approach. By overcoming these challenges, the λ parameter in this proposed model would tend towards one, and according to Equation (20), the variance of observed flows would decrease with no disturbance in the network.

As to this proposed model, let S be the set of optimal links obtained from Model 2. The links in S specify the locations where interviews must be run. By conducting O–D interviews from λ percent of the volume passing through the links in S, the O–D matrix $\tilde{T}$ is observed. Each O–D pair may include several paths and the flow from each O–D may be observed in multiple selected links, which in turn raise the following questions: How are the O–D flows generalized? How are the pair flows' SDs calculated?

To respond to these questions with respect to their order assume that O–D pair w is observed on a subset of optimal links Q $(Q \subseteq S)$, therefor, the generalized volume for each O–D pair is calculated through Equation (21):

$$\hat{X}^w = \sum_{a \in Q} m_a^w \cdot \hat{X}_a^w, \quad \forall w \in W. \qquad (21)$$

Moreover, the variance of the flow is obtained through Equation (22):

$$\mathrm{Var}\left(\sum_i a_i \cdot x_i \right) = \sum_i a_i^2 \cdot \mathrm{Var}(x_i) -$$
$$2 \cdot \sum_{i \ne j} a_i \cdot a_j \cdot \mathrm{Cov}(x_i, x_j). \qquad (22)$$

Given the independence assumption between the observed flows of w-th O–D on different links, the O–D flow SD is calculated as follows:

$$SD(X^w) = \sqrt{\sum_{a \in Q} (m_a^w)^2 \cdot \mathrm{Var}(\hat{X}_a^w)}, \quad \forall w \in W. \qquad (23)$$

As obvious in Equation (17) the different O–D flows on a certain link are interrelated.

The proposed models presented above are able to determine O–D matrix uniquely by applying the collected information from the selected links in two situations of where access and no access to path-ID sensors. In the next section, In order to understand the behaviour of these proposed models and methods, the two examples of Nguyen–Dupuis and real network of intercity road network in the Province of Isfahan are selected.

3. Results and discussion

These proposed models are applied to the small Nguyen–Dupuis transportation network and the results are analysed, evaluated, and validated. Furthermore, in order to demonstrate their applicability in real large networks, the results obtained through these models on the intercity road network in the Province of Isfahan are reported as follows.

3.1. Nguyen–Dupuis network

We start with a simple example, so that all the results have a small size to be shown. The Nguyen–Dupuis network is a well-known network with 13 nodes and 38 links, which is frequently applied in transportation researches. The network is illustrated in Figure 2. Table 1 presents the travel demand for each O–D pair in this network. The information pertaining to O–D pairs and the associated demands are extracted from research by Mínguez *et al.* (2010).

3.1.1. Model 1

Assume that the following assumption for Situation 1 in Subsection 2.1 hold true for the Nguyen–Dupuis network: path-ID sensors are applied in the network to collect origin, destination, and path information. Accordingly, to determine the minimum number of path-ID sensors and their locations in the network, Model 1 is applied. Model 1 is a linear binary integer programming model where the
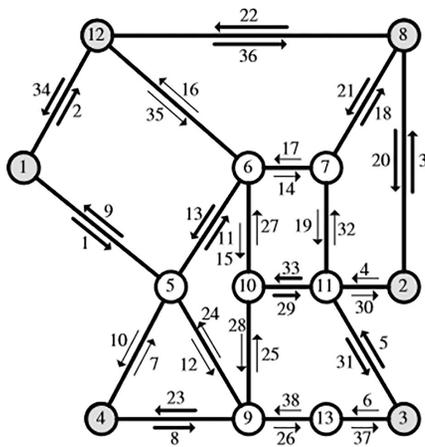
Figure 2. The Nguyen–Dupuis network, showing
the nodes and links

Table 1. O–D pairs and associated demands
in the Nguyen–Dupuis network

No	Origin	Destination	Real demand
1	1	2	2100
2	1	3	4300
3	1	8	3200
4	2	1	2100
5	2	4	3200
6	2	12	500
7	3	1	4300
8	3	4	1100
9	3	12	400
10	4	2	3200
11	4	3	1100
12	4	8	2100
13	8	1	3200
14	8	4	2100
15	8	12	600
16	12	2	500
17	12	3	400
18	12	8	600

Figure 3. Results obtained by applying models
on the Nguyen–Dupuis network: a – results of Model 1;
b – results of Model 3

3.1.2. Model 3

Due to lack of infrastructure and technological development fully equip of networks with path-ID sensors, the interviews are proposed according to the discussion in Subsection 2.3. However, it is extremely difficult if not impossible to record network users' paths (given the number of paths). To overcome this drawback and to determining the number and locations of interviewers in the Nguyen–Dupuis network Model 2 is proposed. This model is an MINLP model, where, the application of which in real large networks is quite challenging and time-consuming. Consequently, Model 2 is rewritten as an MIP model as Model 3 that result of which on the Nguyen–Dupuis network is yield as follows:

$$S_2 = \{6, 14, 17, 22, 26, 33, 36\}, \tag{25}$$

where: S_2 includes the optimal links for conducting interviews in the Nguyen–Dupuis network Model 3. By run interviews on links 6, 14, 17, 22, 26, 29, 33 and 36 and recording only the users' O–D, the Nguyen–Dupuis O–D matrix is determine uniquely. The selected links in this network are distinguished in bold arrows in Figure 3b.

number of constraints and binary variables equal those of the reasonable paths in the network (68 paths in this problem) and the links (38 links in this problem), respectively. Because this problem is linear for real networks, it can be solved in a relatively easy manner in a short time through General Algebraic Modelling System (GAMS) (Rosenthal 2008). The obtained result for Model 1 is as follows:

$$S_1 = \{6, 14, 17, 22, 29, 33, 37\}, \tag{24}$$

where: S_1 is the optimal links for installing path-ID sensors. By installing path-ID sensors on links 6, 14, 17, 22, 29, 33, 36 and 37 and recording network users' paths, the Nguyen–Dupuis O–D matrix is determine uniquely. The selected links in this network are distinguished in bold arrows in Figure 3a.

3.2. Simulator program

After identifying the optimal locations in the network for running interviews, the O–D matrix is generalized and evaluated base on the collected information. Here, to evaluate the results, a program that simulating the conditions of the situation is implemented on a computer. The program takes the real O–D matrix together with θ and λ as the inputs to generate sample O–D trip information and selected links volume. In the simulator program, the real O–D matrix is assigned to the network by applying Dial's traffic assignment method with θ parameter, where the links' volumes are obtained. Thereafter, a proportion λ of the traffic passing through each link in set S is randomly selected and the information of O–Ds in this set are determined. In this manner, the O–D information gathered by interviewers are simulated (Figure 4).

By applying this information obtained from this simulated program, the O–D matrix $\left(\tilde{T}\right)$ is obtained which is then applied in determining the generalized O–D matrix and the corresponding SD through Equations (21) and (23).

3.3. Evaluating the generalized O–D matrix

In order to evaluate this matrix, the simulator program is run many times with different θ and λ values. The generalized O–D matrix values against values in the real O–D matrix, where θ = 0.5, λ = 0.1, number of iterations = 50, are plotted in Figure 5.

A careful observation at the horizontal axis in Figure 5 reveals that there exist 7 points rather than 18 on the axis, because some O–D pairs have the same demand values. As observed, due to the fact that only a proportion of the network users are interviewed (λ = 0.1), the generalized demand values are somewhat different from real demand values, containing some error, which can be reduced by in-

creasing λ, as shown in Figure 6. The diagrams in this figure exhibit that relative error of the generalized O–D against real demand $\left(REG = \dfrac{\hat{T} - T_{real}}{T_{real}} \cdot 100 \right)$ for different λ values and 50 iterations of the simulator program at θ = 0.5. As observed in Figure 6, higher λ values decrease the dispersion of relative error values.

3.4. Intercity road network in the Province of Isfahan

Here this proposed model is applied to a real large transportation network, Isfahan's intercity road network in Iran. The Province of Isfahan covers 107027 km² constituting 6.49% of Iran. According to the 2011 census, the province has a population of 4879312, among which 1908968 dwell the City of Isfahan. The transportation network in the province consists of 248 links and 77 nodes. 24 of these nodes are deemed as trip generation (origin) or trip attraction (destination) nodes, leading to the dimensions of O–D matrix with 22×22. The locations of O–Ds in Isfahan's intercity road network are shown in Figure 7.
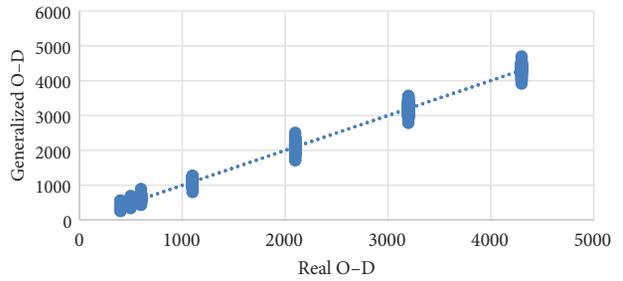
Figure 5. Generalized O–D matrix versus real O–D matrix for the Nguyen–Dupuis matrix
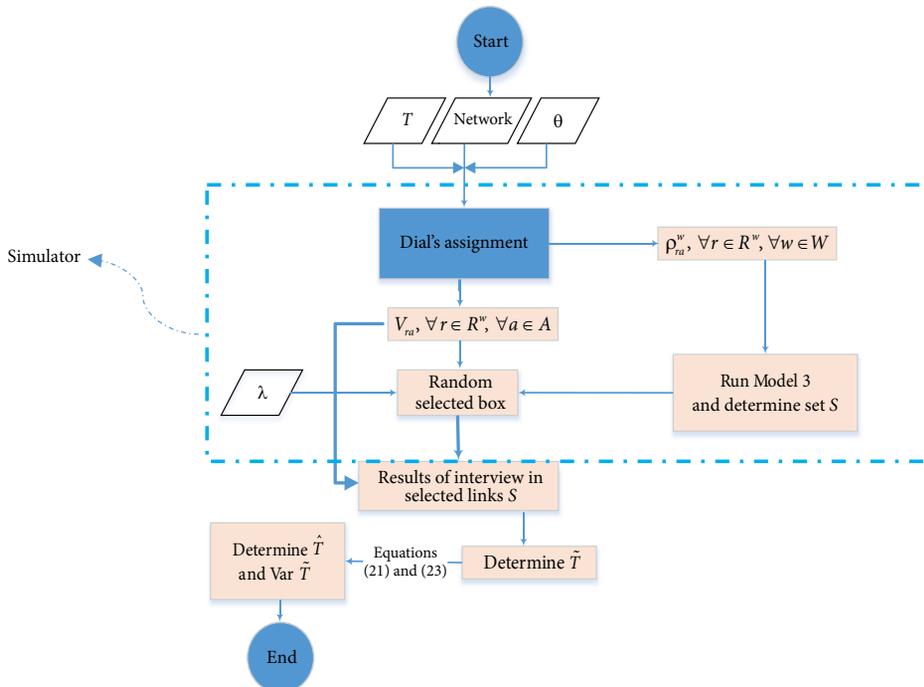
Figure 4. Flowchart of the proposed methodology
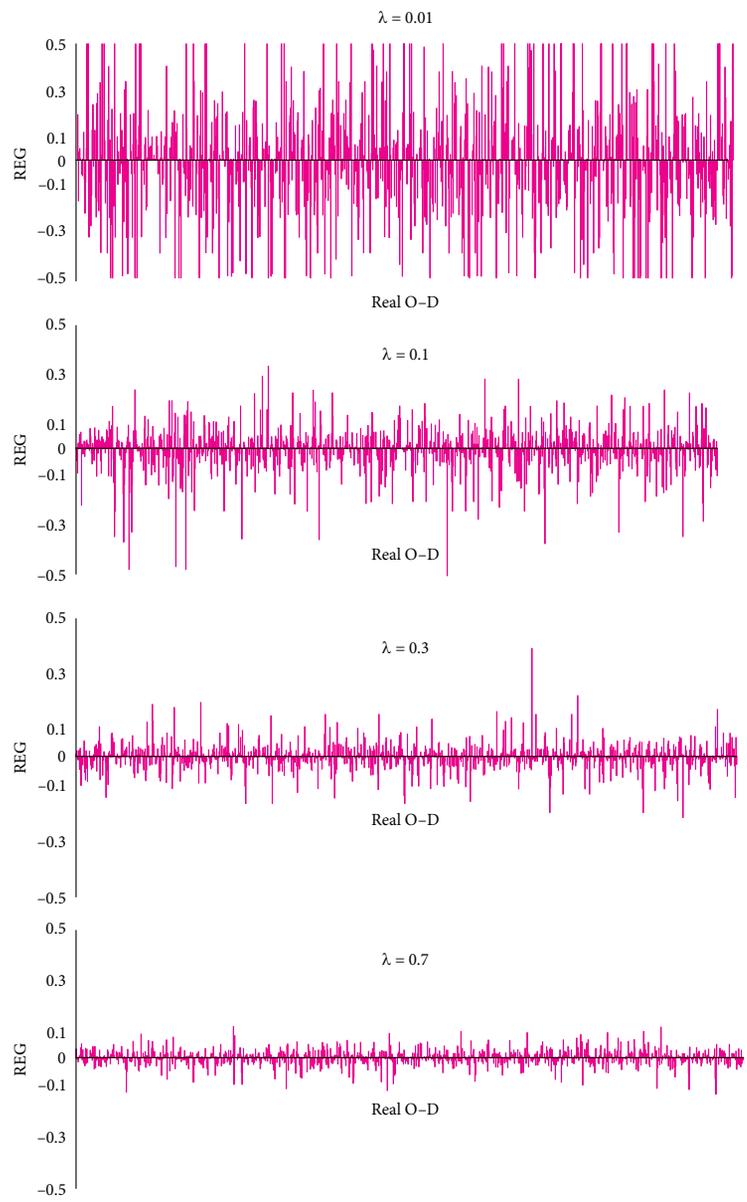
Figure 6. Relative error of the generalized O–D relative versus real demand for different values of λ
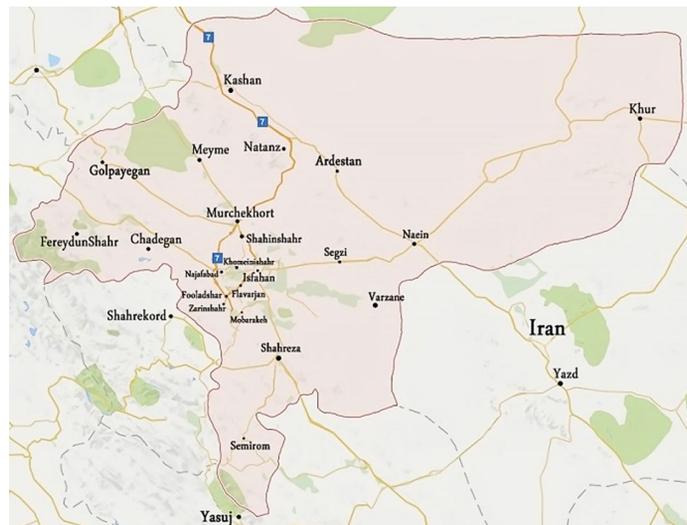
Figure 7. Locations of O–Ds in Isfahan's intercity road network
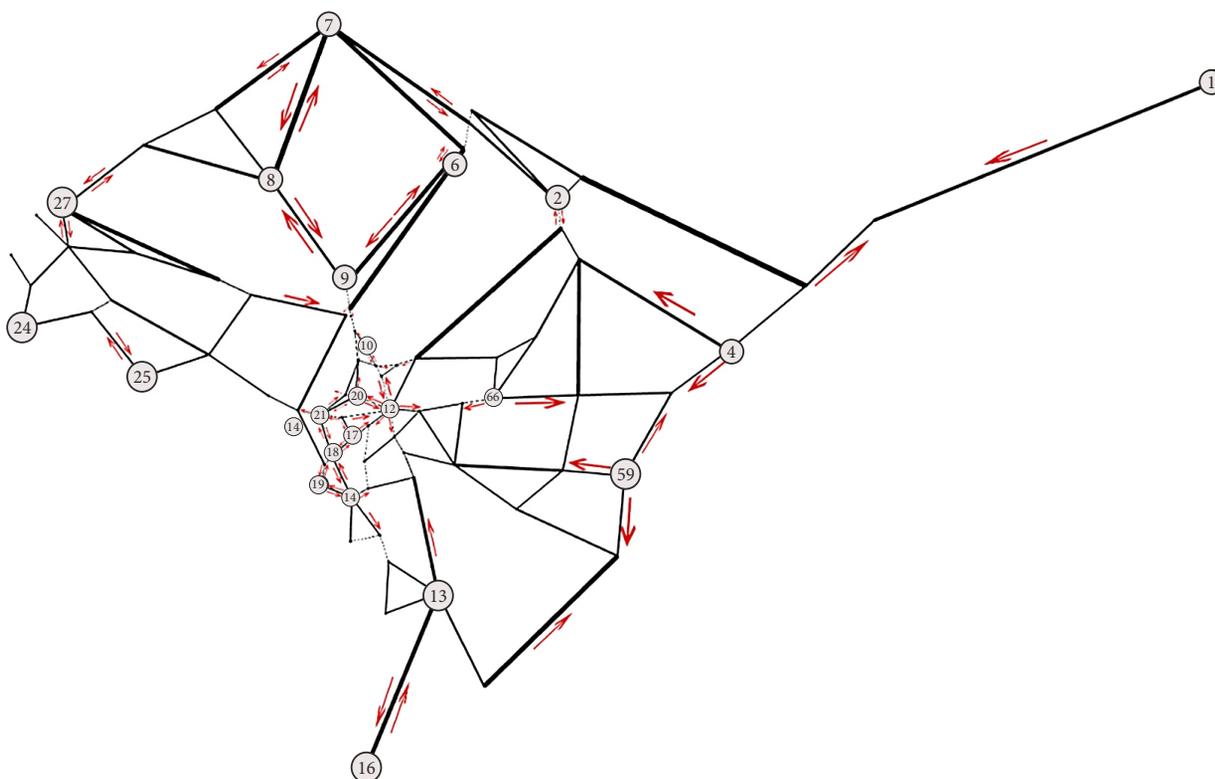
Figure 8. Isfahan's intercity road network marking the links selected for conducting interviews

As observed, Isfahan's intercity road network is a real, uncongested, and relatively large network and the assumptions in Model 3 hold true for this network. This model is solved in GAMS through a PC Computer with 4 GB RAM and 2.6 GB Core i7 CPU in 10.156 s where 73 (set S_3) of the 248 links (nearly 30% of the links) are selected. The locations of these links for running interviews are exhibited in arrows in Figure 8.

$$S_3 = \{3, 4, 6, 7, 11, 12, 15, 16, 17, 18, 19, 20, 21,$$
$$23, 25, 28, 31, 32, 33, 34, 35, 36, 39, 41, 43, 44,$$
$$45, 46, 49, 50, 51, 54, 55, 59, 60, 63, 64, 78, 92,$$
$$99, 100, 101, 104, 106, 111, 112, 116, 118, 119,$$
$$120, 121, 124, 128, 141, 155, 157, 159, 168, 181,$$
$$182, 193, 195, 209, 216, 219, 220, 221,$$
$$222, 229, 234, 236, 245, 246\}. \qquad (26)$$

Conclusions

One of the most important problems in transportation planning is identifying the accurate O–D matrix. O–D matrix observation is a problem where the flow information gathered from network leads to a unique solution for the O–D matrix. Observing this matrix heavily depends on how to collect path flows' data from network. In recent years, with the advancement of technology, a generation of sensors are developed that can collect path flows in the transportation networks. Path-ID sensors are of the type,

which are capable of identifying network users' paths. Unlike other available sensors, path-ID sensors have not gained enough interest among researches. In this paper, the problem of finding the location and the minimum number of path-ID sensors to observe O–D matrix is solved.

Nowadays with advent of new technologies like cell phones and in-vehicle GPSs it is possible to collect accurate and reliable information at low cost, while due to the low penetration rates of these devices among transportation networks' users, excessive costs and the probability of statistical bias for specific social groups (such as young people or business people), a lot is yet to be done to extract sensible and valid information from these new data sources. However, combining traditional survey data with advanced data sources (e.g. AVI, GPS, and CDRs) provide consistency in recording long-term travel information. Here, it is rational to deduce that, in near future, low cost data collection will be feasible in real-time mobility traces from multiple and heterogeneous sources like buses, taxis, trams, private cars, individual smartphones, etc. in order to monitor traffic in a rapid, accurate, economic and continuous basis.

With respect to the restrictions available in advanced data sources, the idea and functionality of them are applied through the run interviews. Based on which some mathematical programming models are proposed to determine O–D matrix uniquely. Here, with respect to functional and budgetary constraints together with the objec-

tive of not disrupting the network, the interview is run on a percentage of users. The data obtained from the interview are generalized and the variance of these flows are calculated. The generalized flows, their variances and use of another data sources, like traffic counts, allow variances to be reduced and accuracy to be increased (e.g. Castillo *et al.* 2013).

To illustrate the operational capability of the above mentioned models in practice, they are implemented on a large real network, the intercity road network in the Province of Isfahan (Iran). This case study demonstrates that these proposed approaches are applicable in large networks.

Because the objective of this paper is full O–D matrix observation with no initial data, total sensor cost (device, installation, and maintenance) is not addressed here. On the contrary, when the objective is partial O–D matrix observation oriented, different constraints like sensor installation cost and location-specific restrictions become of concern. Therefore, a trade-off between amount of measured/observed information and solution feasibility is established.

Future research steps should include research along three main directions. The first one is full O–D matrix observation through combining information collected from various types of sensors (both active and passive). The second direction of further research could focus on the observation of priority O–D pairs flow and considering sensor installation costs. A third direction of further research includes finding the minimum number and locations of new sensors in accordance with the existing ones in order to observe the O–D matrix.

References

Bauer, D.; Richter, G.; Asamer, J.; Heilmann, B.; Lenz, G.; Kölbl, R. 2018. Quasi-dynamic estimation of OD flows from traffic counts without prior OD matrix, *IEEE Transactions on Intelligent Transportation Systems* 19(6): 2025–2034. https://doi.org/10.1109/TITS.2017.2741528

Caceres, N.; Wideberg, J. P.; Benitez, F. G. 2008. Review of traffic data estimations extracted from cellular networks, *IET Intelligent Transport Systems* 2(3): 179–192. https://doi.org/10.1049/iet-its:20080003

Castillo, E.; Calviño, A.; Lo, H. K.; Menéndez, J. M.; Grande, Z. 2014. Non-planar hole-generated networks and link flow observability based on link counters, *Transportation Research Part B: Methodological* 68: 239–261. https://doi.org/10.1016/j.trb.2014.06.015

Castillo, E.; Cobo, A.; Jubete, F.; Pruneda, R. E.; Castillo, C. 2001. An orthogonally based pivoting transformation of matrices and some applications, *SIAM Journal on Matrix Analysis and Applications* 22(3): 666–681. https://doi.org/10.1137/S0895479898349720

Castillo, E.; Nogal, M.; Rivas, A.; Sánchez-Cambronero, S. 2013. Observability of traffic networks. Optimal location of counting and scanning devices, *Transportmetrica B: Transport Dynamics* 1(1): 68–102. https://doi.org/10.1080/21680566.2013.780987

Castillo, E.; Rivas, A.; Jiménez, P.; Menéndez, J. M. 2012. Observability in traffic networks. Plate scanning added by counting information, *Transportation* 39(6): 1301–1333. https://doi.org/10.1007/s11116-012-9390-0

Chootinan, P.; Chen, A.; Yang, H. 2005. A bi-objective traffic counting location problem for origin-destination trip table estimation, *Transportmetrica* 1(1): 65–80. https://doi.org/10.1080/18128600508685639

De Grange, L.; González, F.; Bekhor, S. 2017. Path flow and trip matrix estimation using link flow density, *Networks and Spatial Economics* 17(1): 173–195. https://doi.org/10.1007/s11067-016-9322-1

Földes, D.; Csiszár, C. 2015. Route plan evaluation method for personalised passenger information service, *Transport* 30(3): 273–285. https://doi.org/10.3846/16484142.2015.1086889

Gentili, M.; Mirchandani, P. B. 2012. Locating sensors on traffic networks: Models, challenges and research opportunities, *Transportation Research Part C: Emerging Technologies* 24: 227–255. https://doi.org/10.1016/j.trc.2012.01.004

He, S.-X. 2013. A graphical approach to identify sensor locations for link flow inference, *Transportation Research Part B: Methodological* 51: 65–76. https://doi.org/10.1016/j.trb.2013.02.006

Hu, S.-R.; Peeta, S.; Chu, C.-H. 2009. Identification of vehicle sensor locations for link-based network traffic applications, *Transportation Research Part B: Methodological* 43(8–9): 873-894. https://doi.org/10.1016/j.trb.2009.02.008

Kim, H.; Nam, D.; Suh, W.; Cheon, S. H. 2018. Origin-destination trip table estimation based on subarea network OD flow and vehicle trajectory data, *Transportation Planning and Technology* 41(3): 265–285. https://doi.org/10.1080/03081060.2018.1435437

Lee, R. J.; Sener, I. N.; Mullins, J. A. 2016. An evaluation of emerging data collection technologies for travel demand modeling: from research to practice, *Transportation Letters: the International Journal of Transportation Research* 8(4): 181–193. https://doi.org/10.1080/19427867.2015.1106787

Li, X.; Kurths, J.; Gao, C.; Zhang, J.; Wang, Z.; Zhang, Z. 2018. A hybrid algorithm for estimating origin-destination flows, *IEEE Access* 6: 677–687. https://doi.org/10.1109/ACCESS.2017.2774449

Mínguez, R.; Sánchez-Cambronero, S.; Castillo, E.; Jiménez, P. 2010. Optimal traffic plate scanning location for OD trip matrix and route estimation in road networks, *Transportation Research Part B: Methodological* 44(2): 282–298. https://doi.org/10.1016/j.trb.2009.07.008

Mitsakis, E.; Chrysohoou, E.; Salanova Grau, J. M.; Iordanopoulos, P.; Aifadopoulou, G. 2017. The sensor location problem: methodological approach and application, *Transport* 32(2): 113–119. https://doi.org/10.3846/16484142.2016.1258674

Ng, M. 2012. Synergistic sensor location for link flow inference without path enumeration: A node-based approach, *Transportation Research Part B: Methodological* 46(6): 781–788. https://doi.org/10.1016/j.trb.2012.02.001

Olia, A.; Abdelgawad, H.; Abdulhai, B.; Razavi, S. N. 2017. Optimizing the number and locations of freeway roadside equipment units for travel time estimation in a connected vehicle environment, *Journal of Intelligent Transportation Systems* 21(4): 296–309. https://doi.org/10.1080/15472450.2017.1332524

Ortúzar, J. de D.; Willumsen, L. G. 2011. Trip distribution modelling, in J. de D. Ortúzar, L. G. Willumsen (Eds.). *Modelling Transport*, 175–206. https://doi.org/10.1002/9781119993308.ch5

Pravinvongvuth, S. 2007. *Two Location Problems in Transportation*. PhD Thesis. Utah State University, US. 220 p.

Rinaldi, M.; Viti, F. 2017. Exact and approximate route set generation for resilient partial observability in sensor location problems, *Transportation Research Part B: Methodological* 105: 86–119. https://doi.org/10.1016/j.trb.2017.08.007

Rosenthal, R. E. 2008. *GAMS – a User's Guide*. GAMS Development Corporation, Washington, DC, US. 293 p.

Schrijver, A. 1998. *Theory of Linear and Integer Programming*. Wiley. 484 p.

Viti, F.; Cantelmo, G.; Corman, F.; Rinaldi, M. 2015. Improving the reliability of demand estimation using traffic counts by including information on link flow observability, in *6th International Symposium on Transportation Network Reliability* (*INSTR 2015*), 2–3 August 2015, Nara, Japan. https://doi.org/10.3929/ethz-b-000183287

Xu, X.; Lo, H. K.; Chen, A.; Castillo, E. 2016. Robust network sensor location for complete link flow observability under uncertainty, *Transportation Research Part B: Methodological* 88: 1–20. https://doi.org/10.1016/j.trb.2016.03.006