Transportation Research Part B 89 (2016) 82–106

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Vehicle identification sensor models for origin–destination estimation

Majid Hadavi[a,*], Yousef Shafahi[b,1]

[a] Sharif University of Technology, Azadi St, Tehran, Iran
[b] Department of Civil Engineering, Sharif University of Technology, Azadi Ave., 11155-9313 Tehran, Iran

ARTICLE INFO

Article history:
Received 22 April 2015
Revised 27 February 2016
Accepted 30 March 2016
Available online 19 April 2016

Keywords:
OD estimation
Location model
Vehicle identification sensors

ABSTRACT

The traditional approach to origin–destination (OD) estimation based on data surveys is highly expensive. Therefore, researchers have attempted to develop reasonable low-cost approaches to estimating the OD vector, such as OD estimation based on traffic sensor data. In this estimation approach, the location problem for the sensors is critical. One type of sensor that can be used for this purpose, on which this paper focuses, is vehicle identification sensors. The information collected by these sensors that can be employed for OD estimation is discussed in this paper. We use data gathered by vehicle identification sensors that include an ID for each vehicle and the time at which the sensor detected it. Based on these data, the subset of sensors that detected a given vehicle and the order in which they detected it are available. In this paper, four location models are proposed, all of which consider the order of the sensors. The first model always yields the minimum number of sensors to ensure the uniqueness of path flows. The second model yields the maximum number of uniquely observed paths given a budget constraint on the sensors. The third model always yields the minimum number of sensors to ensure the uniqueness of OD flows. Finally, the fourth model yields the maximum number of uniquely observed OD flows given a budget constraint on the sensors. For several numerical examples, these four models were solved using the GAMS software. These numerical examples include several medium-sized examples, including an example of a real-world large-scale transportation network in Mashhad.

1. Introduction

Origin–destination (OD) estimation based on exhaustive data surveys can be replaced with reasonable low-cost approaches such as OD estimation based on the data collected by traffic sensors (Doblas and Benitez, 2005). In OD estimation based on traffic flow observations, the problem of the locations of the sensors is critical. This location problem involves the selection of certain arcs or nodes for the sensors. These models minimize the number of sensors and maximize the precision and reliability of the final estimated target flows. Here, the term target flows refers to the flows that need to be estimated. Instead of the OD vector, these target flows may also be path flows or link flows (Castillo et al., 2008, 2013; Gentili and Mirchandani, 2012; Ng, 2012). Gentili and Mirchandani (2012) reviewed the literature and presented a "unifying picture" of sensor location models. They divided these location models into two major categories:

* Corresponding author at: Third Ring, Door Number 5, Mirzaie Avenue, Ajanlou St, Mojahedin Eslam St, Tehran 1157748763, Iran. Tel.: +98 9128442915. *E-mail addresses:* mjdhdv@gmail.com, majid_hadavi2000@yahoo.com (M. Hadavi), shafahi@sharif.edu (Y. Shafahi).
[1] Tel.: +98 21 66164246; fax: +98 21 66014828.

http://dx.doi.org/10.1016/j.trb.2016.03.011
0191-2615/

- Sensor location flow-observability models: These models establish certain conditions for the locations of the sensors such that the target flows are unique.
- Sensor location flow-estimation models: These models establish the conditions for the locations of the sensors that yield the best estimation of the target flows.

Gentili and Mirchandani (2012) also identified four categories of sensor types: counting sensors, path-ID sensors, image sensors and vehicle identification sensors. Such sensors have also been used in various studies for purposes other than flow estimation, such as travel time measurement. Some of these studies have used counting sensors (Viti et al., 2008; Li and Ouyang, 2011), and others have used vehicle identification sensors (Sherali et al., 2006; Mirchandani et al., 2009). However, the present paper is focused on the estimation of target flows.

The literature on flow estimation based on sensor observations can be divided into twelve categories based on the type of sensors used (four types) and the target flows considered (three types). Certain of these twelve categories have received more attention than others:

- Estimating OD flows based on counting sensor observations (Yang et al., 1991; Yang and Zhou, 1998; Doblas and Benitez, 2005; Gan et al., 2005; Ehlert et al., 2006; Eisenman et al., 2006; Viti et al., 2014)
- Estimating arc flows based on counting sensor observations (Hu et al., 2009; Ng 2012, 2013; Castillo et al., 2013)
- Estimating arc flows based on image sensor observations (Bianco et al., 2006)
- Estimating path flows based on path-ID sensor observations (Gentili and Mirchandani, 2005)
- Estimating OD flows based on vehicle identification sensor observations (Castillo et al., 2008; Mínguez et al., 2010; Zhou and List, 2010)
- Estimating path flows based on vehicle identification sensor observations (Castillo et al., 2008, 2010; Mínguez et al., 2010).

Several papers have also discussed the simultaneous use of both counting sensors and vehicle identification sensors to estimate target flows and the corresponding location problem (Zhou and List, 2010). The current paper focuses on the locations of vehicle identification sensors considering two types of target flows: OD flows and path flows.

Viti et al. (2014) have presented an "extensive review" of location models for traffic sensors. Their review is more recent than that of Gentili and Mirchandani (2012) and covers several additional newer papers. In what follows, some of the literature on flow estimation based on counting sensors and vehicle identification sensors will be reviewed.

1.1. Counting sensors

For the case in which counting sensors are used and the target flow is the OD vector, Yang et al. (1991) proposed the use of the maximum possible relative error, or MPRE. The MPRE is an index of the precision or reliability of an OD estimation, with a reasonable theoretical background. This index can be employed either for determining sensor locations or for OD estimation. Yang and Zhou (1998) formulated four rules with regard to the MPRE index. According to these rules, the placement of traffic counting sensors should ensure the following conditions: 1- For each OD pair, a sensor should be located on at least one significant arc (an arc that contains more than a certain portion of an OD pair flow). 2- For each OD pair, the portion of the demand that is covered should be maximized. 3- The observed net flow (the sum of the flows on paths with at least one sensor) should be maximized. 4- The traffic counts on the observed links should be linearly independent of each other.

Gan et al. (2005) proposed the use of the expected relative error, or ERE, based on the MPRE. Whereas the MPRE measures the quality of an OD estimation in terms of the maximum possible relative error, an alternative means of quantifying the quality of an OD estimation is to measure the expected value of its relative error. In other words, the MPRE is the distance between the estimated OD vector and the most distant feasible OD vector, whereas the ERE is the expected distance between the estimated OD vector and a random feasible OD vector. Several other measures of the quality of an OD estimation also exist, such as the total demand scale, or TDS (Bierlaire, 2002), and the relative total demand scale, or RTDS (Gan et al., 2005).

Ehlert et al. (2006) proposed a linear integer model for determining the locations of counting sensors. Their model considers an index of importance for each OD pair and maximizes the sum of the importance indices of the covered OD pairs in the objective function. In their study, an OD pair was considered to be covered if a sensor was located on at least one significant arc for that OD pair. According to Ehlert et al. (2006), this model can be solved for medium and large networks because of the reasonable upper bounds provided by the LP sub-problem. Ehlert et al. (2006) solved their model for networks with more than 1000 links within a short period of time.

Hu et al. (2009) addressed the full observability problem for link flows based on link counting sensors from a "budgetary planning perspective". They divided the set of links into "base links" and "non-base links". The base links, according to their definition, constitute the smallest set of links for which counting their flows results in unique flows through all links. These authors proposed a Gaussian elimination algorithm to identify the base links.

Ng (2012) reformulated the link observability problem proposed by Hu et al. (2009). In his reformulation, path enumeration and the construction of the complete arc-route incidence matrix are no longer needed. This makes his approach more appropriate for and applicable to large networks. He presented an upper bound on the number of sensors required for full observability of the link flows. In another paper (Ng, 2013), Ng proposed an algorithm for solving a more general problem to answer the following question: "Suppose that one is only interested in a subset of link flows and that the link flows are known a priori. At a minimum, what link flows are needed to be able to uniquely determine the desired link flows?"

Castillo et al. (2013) showed that when path information is used, Ng's upper bound can be reduced. They also showed that this reduction can be as large as 16% and claimed that it can be greater for larger networks. They suggested the use of a set of linearly independent paths instead of the entire set of paths. They also presented a computationally efficient algorithm for obtaining a minimum set of linearly independent paths.

Viti et al. (2014) presented an "intuitive metric" to evaluate a given set of sensor locations. In that paper, the focus was placed on counting sensors. The presented metric is useful in real networks in which it is not economically feasible to achieve full observability of the target flows. Based on this metric, Viti et al. (2014) proposed a greedy heuristic algorithm to find the links that contain the most information. They defined information acquisition as "the process by which the size of the solution space is reduced, either by measuring some of the variables or exploiting existing relations between them".

1.2. Vehicle identification sensors

As mentioned previously, this paper focuses on path flow and OD vector estimation based on vehicle identification sensors. Castillo et al. (2008) proposed a method for estimating path flows based on the data provided by vehicle identification sensors. They also described the format in which these sensors provide observations of path flows. This format does not include the order in which the vehicles were detected by different sensors. More precisely, Castillo et al. (2008) assigned a subset of the sensors to each detected vehicle. Based on this information format, they proposed a location model that yields the minimum number of sensors required for full observability of all considered path flows.

The data provided by vehicle identification sensors have also been discussed by Gentili and Mirchandani (2012). In the current paper, the format in which the data are presented is the same as that of Gentili and Mirchandani and differs from that of Castillo et al. (2008, 2010). Castillo et al. (2008) considered only unordered subsets of sensors to constitute the combinations of sensors for analysis; however, because the vehicle ID detection data have time stamps, the order of detection of these sensors is also available. The order of the sensors is included in the format presented by Gentili and Mirchandani (2012). Considering the order of the sensors can reduce the minimum number of sensors required to ensure full observability of the path flows. This is demonstrated in the current paper by means of an example.

For the remainder of this paper, we assume that the orders of the sensors by which vehicles are detected are available in the data provided by vehicle identification sensors. Based on this assumption, a new location model for the uniqueness of path flows is proposed. In Theorem 2, we claim and prove that in this case,[2] our location model always yields the minimum number of sensors to ensure the uniqueness of path flows. To our knowledge, no other location models exist that consider the order of the sensors.

Mínguez et al. (2010) considered the location problem for vehicle identification sensors. They proposed a new location model based on the model proposed by Castillo et al. (2008). In their model, a budget constraint is included. The incorporation of this constraint renders the model more applicable. At the time it was proposed, the model of Mínguez et al. (2010) was the only model available for determining the locations of vehicle identification sensors while considering a budget constraint. Because no other such models were available, they compared their model with models for counting sensors. By running simulations, they compared their location model with a number of existing models for the selection of the locations of counting sensors. Thus, they demonstrated that their model yields superior results.

In addition to discussing the location model proposed by Castillo et al. (2008) and proposing a more complete model, we also propose several alternative location models. In the model of Castillo et al. (2008) and our first model, L_1, the objective is to uniquely observe all path flows. In large networks, observing all path flows is extremely expensive. However, many reasonable methods are available for assigning an OD vector to the network and thus obtaining the path flows of interest. We therefore modify the location model of Castillo et al. (2008) and our first model (the modified version is called L_3) for the unique observation of the OD vector; this approach can significantly reduce the number of sensors required. We claim and prove that our modified model, L_3, always yields the minimum number of sensors required to uniquely observe the OD vector. Moreover, two other models are presented in this paper. The first, L_2, maximizes the number of fully observable path flows subject to a budget constraint. The second, L_4, maximizes the number of fully observable OD flows subject to a budget constraint. These four models have been solved for four examples using the topologies of the Nguyen–Dupuis network, the Sioux–Falls network and the Mashhad network. The results are reported in Section 4.

This paper is organized as follows: Section 2 introduces vehicle identification sensors in terms of the observations they provide. In Section 3, some of the existing models are discussed and several new models are proposed. Section 3 consists of two subsections. Section 3.1 addresses the location models in which the objective is to uniquely observe all path flows. Section 3.2 addresses the location models in which the objective is to uniquely observe the OD vector. In Section 4, four examples are solved for the models proposed in Sections 3.1 and 3.2. Section 5 presents the conclusions.

[2] In the case that the sensor order is considered.

2. Vehicle identification sensors

A traditional counting sensor provides only the volume that corresponds to the arc on which it is located. However vehicle identification sensors provide not only the volume of the arc on which they are located but also the volumes that correspond to different combinations of sensors. For example, if the three arcs A, B and C in a network each have an ALPR camera (Automatic License Plate Recognition camera), then the volumes of these arcs can be observed. The number of vehicles that pass arc A and then pass arc B but do not pass arc C can be observed. Similarly, the number of vehicles that pass arc B and then pass arc A but do not pass arc C can also be observed. Generally, the volumes corresponding to all non-empty combinations of sensors can be observed: A, B, C, AB, BA, AC, CA, BC, CB, ABC, ACB, BAC, BCA, CAB, and CBA. Although the majority of these combinations are useless in practice and their corresponding volumes are zero, the number of combinations exceeds the number of sensors. Therefore, vehicle identification sensor observations can be extremely informative compared with those of traditional counting sensors, especially in large networks.

In the data provided by vehicle identification sensors, or simply sensors,[3] three attributes are available for each vehicle detection event (Castillo et al., 2008; Gentili and Mirchandani, 2012):

- An identification code for the vehicle
- The time at which the sensor observed the vehicle
- The arc on which that sensor is installed

The vehicle identification code may be the license plate number, as in the case of ALPR cameras. Among all such records over a given study period, each detected identification code (in other words, each detected vehicle) will have been recorded by one or more sensors. Consider, for example, a vehicle that has been recorded once by each of the vehicle identification sensors on arcs 5, 12, 32, 93 and 102, where the corresponding times for each detection event are 7:52 AM, 7:00 AM, 6:23 AM, 6:43 AM, and 7:12 AM, respectively (assume, for example, that the given study period is from 5:00 AM to 9:00 AM on a specific day). Taken together, these records show that the corresponding vehicle passed arcs 5, 12, 32, 93 and 102 in the following order: 32–93–12–102–5. Thus, we say that this vehicle corresponds to the combination (32, 93, 12, 102, 5), where the term combination means an ordered subset of sensors. The total number of detected vehicles in the entire database that correspond to the combination (32, 93, 12, 102, 5) is called the corresponding volume of this combination. Each non-empty combination has a corresponding volume that is known from the sensor observations.

The corresponding volume of each combination is the sum of the flows on the subset of paths that matches that combination. These are called the corresponding paths of that combination. This subset comprises the set of paths that contain the arcs corresponding to the sensors in this combination in the same order in which they appear in the combination. For example, the corresponding paths of the combination (32, 93, 12, 102, 5) are the paths that contain arcs 32, 93, 12, 102 and 5, in that order. Thus, each combination of sensors,[4] or combination, has a corresponding volume and corresponds to a subset of network paths.

A distinct but important feature of different combinations of sensors is that each path corresponds to one combination. The set of paths that correspond to a given combination has no common element with the corresponding paths of another combination. When the locations of the sensors are fixed, the set of sensored arcs that belong to each path and the order of those arcs are uniquely defined; thus, one path cannot correspond to more than one combination. Therefore,

$$R'_j \cap R'_{j'} = \emptyset \quad \forall j \neq j' \in C \tag{1}$$

where C and R'_j are the set of all possible combinations of sensors and the set of paths that correspond to sensor combination $j \in C$, respectively. As mentioned earlier in this section, each non-empty combination has a corresponding volume, which can be zero. This volume is equal to the sum of the flows on its corresponding paths. Thus, the equations obtained from the sensor observations are as follows:

$$\sum_{r \in R'_j} f_r = b_j \quad \forall j \in C \mid j \neq \emptyset \tag{2a}$$

where f_r and b_j are the flow on path "r" and the corresponding volume of sensor combination $j \in C$, respectively. The feasibility of flow or conservation of flow condition is as follows:

$$\sum_{r \in R_i} f_r = Q_i \quad \forall i \in I, \;\; f_r \geq 0 \quad \forall r \in R \tag{2b}$$

where R_i, Q_i and I are the set of paths that correspond to OD pair "i", the volume of OD pair "i" and the set of all OD pairs, respectively. Once the observations from the sensors over a defined period of study have been gathered, the vector of the path flows and the OD vector must satisfy these two constraints. We will henceforth refer to these two constraints as condition (2).

[3] For brevity, we will use the term sensors to refer to vehicle identification sensors throughout the remainder of the paper.

[4] By a combination of sensors, we refer to an ordered subset of sensors.

3. The location problem for vehicle identification sensors

3.1. Sensor location problem for the uniqueness of path flows

This section primarily discusses the establishment of conditions for the locations of the sensors such that the path flows are unique. For example, if a sensor is located on every arc in a network, then condition (2) will have only one unique solution in terms of f and Q.

The flows observed by the sensors are attributed to an unknown but fixed demand vector. The purpose of placing these sensors is to estimate this vector. Let us assume that the observations of the sensors are free of error. In this case, the set of demand vectors that satisfy the observation constraint (2a) and the feasibility constraint (2b) contains that true demand vector and thus is always non-empty. Therefore, in this study, the set of path flows and OD vectors that satisfy condition (2) is assumed to always be non-empty; this result is independent of the sensor locations and the number of sensors, which are defined by vector J.

3.1.1. The existing model

Castillo et al. (2008) proposed a location model for vehicle identification sensors to determine the minimum number of sensors required for unique observation of the path flows. This model is formulated as follows:

$$C: \ \ Minimize \ n_s = \Sigma_{a\in A} J_a$$

$$J_a \in \{0, 1\} \qquad \forall a \in A \tag{3}$$

$$\sum_{a\in A} J_a d(r_1, r_2, a) \geq 1 \qquad \forall r_1, r_2 \in R | r_1 \neq r_2 \tag{4}$$

$$\sum_{a\in A} J_a \delta_{ra} \geq 1 \qquad \forall r \in R \tag{5}$$

where J_a is a binary variable whose value is one when arc "a" has a sensor and zero otherwise. A is the set of all arcs in the network. δ_{ra} is a binary parameter whose value is one when path "r" contains arc "a" and zero otherwise. $d(r_1, r_2, a)$ is a binary parameter whose value is one if one of the paths r_1 and r_2 contains arc a and the other path does not contain arc a and whose value is zero otherwise.

Castillo et al. (2008) claimed that the optimal solution to model C contains the least number of sensors required to uniquely observe all path flows. In this model, the order in which each vehicle is detected by the sensors is not considered. However, as mentioned previously, because the vehicle ID detection data have time stamps, the order of these detections is also available from the sensor observations. Model C was proposed based on the assumption that the sensor order is not available. Under this assumption, model C always yields the minimum number of sensors required for full observability of the selected path flows. By contrast, in this study, we adopt the assumption that the order of the sensors is available. Here, we explore the performance of model C in this new setting. As mentioned before, we assume that the observations of the sensors are error-free and conclude that the set of path flows that satisfy condition (2) is always non-empty. A theorem regarding this model is presented here:

Theorem 1. *Assume that J is the sensor location vector and that the order in which the vehicles are detected by the sensors is available. The constraints of model C are sufficient but not necessary to ensure the uniqueness of the path flows that satisfy condition (2).*

The proof of sufficiency is similar to the proof of Theorem 3 that is presented in Section 3.2. The following example shows that these constraints are no longer necessary. In this example, we identify a sensor location vector that does not satisfy constraint (4) but ensures the uniqueness of the path flows.

Example 1. Consider the network illustrated in Fig. 1.

Consider two OD pairs for this network: from 1 to 3 and from 2 to 4. The first OD pair corresponds to two possible paths: {1–2} and {6–3–2}. The second OD pair also corresponds to two paths: {5–6} and {2–4–6}. Assume a location vector for the sensors in which only arcs 2 and 6 are equipped with a sensor. We claim that this location vector will result in unique path flows but does not satisfy constraint (4) of model C.

For this sensor location vector, the combinations corresponding to each of the paths identified above are (2), ((6–2), (6) and (2–6), respectively. Thus, each path corresponds to a different non-empty combination, and the path flows can be uniquely defined based on the sensor observations. The reason for this uniqueness can be seen when checking constraint (2a) for this scenario:

$$\sum_{r\in R'_{(2)}} f_r = f_1 = b_{(2)}$$

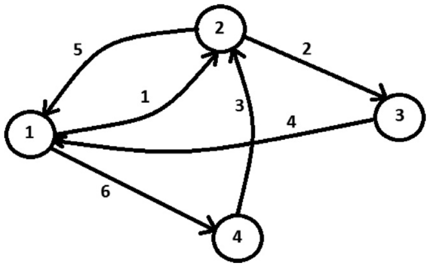

Fig. 1. Network considered in Example 1.

Table 1
Numbers of combinations and subsets of sensors that can be generated for different numbers of sensors.

Number of sensors	10	50	100	150
Number of subsets	1.0e3	1.1e15	1.3e30	1.4e45
Number of combinations	3.9e9	3.5e79	1.2e188	8.2e307

$$\sum_{r\in R'_{(6-2)}} f_r = f_2 = b_{(6-2)}$$

$$\sum_{r\in R'_{(6)}} f_r = f_3 = b_{(6)}$$

$$\sum_{r\in R'_{(2-6)}} f_r = f_4 = b_{(2-6)}$$

However, this location vector does not satisfy constraint (4). We check this constraint below for the second and fourth paths:

$$\sum_{a\in A} J_a d(2, 4, a) = 0 < 1$$

In this equation, the product of J_a and $d(2, 4, a)$ is summed over all six arcs of the network. The sum is found to be zero, which is not greater than or equal to one.

When the order of the sensors is considered, the location vectors need not satisfy constraints (4) and (5) for the path flows to be unique. Therefore, in this case, these two constraints are not necessary for the uniqueness of the path flows. Another important implication of Example 1 is that when the sensor order is considered, model *C* does not always yield the minimum number of sensors required for unique observation of the path flows. In Example 1, the solution to model *C* is {2, 3, 6}; however, we know that {2, 6} will result in unique path flows and that placing a sensor on arc 3 is unnecessary.

The underlying assumption of model *C* is that two paths correspond to the same combination if and only if their sets of sensored arcs are the same; however, two paths may have the same set of sensored arcs but correspond to different combinations because the orders of these arcs differ. This is the case for the pair consisting of the second and fourth paths in Example 1. In Example 1, the set of sensored arcs is {2, 6} for both the second and fourth paths. However, the order of these two arcs is (6–2) for the second path and (2–6) for the fourth path. Thus, the two paths correspond to different combinations and can be uniquely observed.

3.1.1.1. The importance of considering the sensor order. A promising feature of vehicle identification sensors is their ability to provide numerous equations of flow observations based on different sensor combinations. These combinations can produce many equations via two approaches. The first approach involves generating different combinations by considering different subsets of sensored arcs. The second approach is to generate different combinations by considering different orders for each subset of sensors; however, this possibility has not been employed in any of the existing location models for vehicle identification sensors, including model *C*.

Table 1 presents the numbers of possible combinations for different numbers of sensors when the order of the sensored arcs is and is not considered. As shown in Table 1, considering the order of the sensored arcs can increase the number of different combinations that can be generated. As noted by Castillo et al. (2008), the majority of these combinations do not correspond to any real path. However, increasing the total number of combinations also increases the number of combinations that correspond to at least one path.

Considering the order of the sensored arcs is a difficult task in models such as model *C*, in which the sensor location vector is a decision variable rather than a given parameter. The development of a model for which the solution always contains the least possible number of sensors can be challenging. In Section 3.1.2, such a model is proposed. More precisely, by considering the order of the sensors, a location model is proposed that always yields the minimum number of sensors required to uniquely observe the path flows.

3.1.2. Necessary and sufficient conditions for the uniqueness of path flows

In this section, a model named L_1 is proposed. L_1 is a location model for vehicle identification sensors that considers the order of the sensors and yields the minimum number of sensors required for the uniqueness of the path flows. Before model L_1 is presented, several of its main parameters are explained. As mentioned previously, δ is the path-arc incidence matrix, in which $\delta_{ij} = 1$ if arc j belongs to path i and $\delta_{ij} = 0$ otherwise. Matrix E is a three-dimensional matrix in which E_{rab} is equal to one if arc b follows arc a in path r and is zero otherwise. Based on these two matrices, model L_1 is formulated as follows:

$$L_1: \ Minimize \sum_{a\in A} J_a$$

$$y'_{rr'} \le \sum_{a\in A} J_a d\left(r, r', a\right) \qquad \forall r, r' \in R|\ r \ne r' \tag{6}$$

$$y''_{rr'} \le \sum_{a\in A} \sum_{\substack{a' \in A \\ a' \ne a}} J'_{aa'} \left(E_{raa'} E_{r'a'a} + E_{ra'a} E_{r'aa'}\right) \qquad \forall r, r' \in R|\ r \ne r' \tag{7}$$

$$y_{rr'} \le y'_{rr'} + y''_{rr'} \qquad \forall r, r' \in R|\ r \ne r' \tag{8}$$

$$x_r \le y_{rr'} \qquad \forall r, r' \in R|\ r \ne r' \tag{9}$$

$$x_r \le \sum_{a\in A} \delta_{ar} J_a \qquad \forall r \in R \tag{10}$$

$$x_r = 1 \qquad \forall r \in R \tag{11}$$

$$J'_{aa'} \le J_a \qquad \forall a, a' \in A|\ a \ne a' \tag{12}$$

$$J'_{aa'} \le J_{a'} \qquad \forall a, a' \in A|\ a \ne a' \tag{13}$$

$$J'_{aa'} \ge J_a + J_{a'} - 1 \qquad \forall a, a' \in A|\ a \ne a' \tag{14}$$

$$0 \le x_r,\ J'_{aa'}, y'_{rr'}, y''_{rr'}, y_{rr'} \le 1 \qquad \forall r, r' \in R\Big|\ r \ne r', \quad \forall a, a' \in A\Big|\ a \ne a' \tag{15}$$

$$J_a = 0, 1 \qquad \forall a \in A \tag{16}$$

where J_a is a binary variable and the model is designed such that the values of x_r, $J'_{aa'}$, $y'_{rr'}$, $y''_{rr'}$ and $y_{rr'}$ are all either one or zero. J_a is equal to one if arc a has a sensor and is zero otherwise. $J'_{aa'}$ is equal to one if J_a and $J_{a'}$ are both equal to one and is zero otherwise. x_r is equal to one if the flow of path r can be obtained uniquely and is zero otherwise. $y_{rr'}$ is equal to one if paths r and r' correspond to different combinations of sensors and is zero if both of them correspond to the same combination (the empty set is also a combination). $y'_{rr'}$ is equal to one if there exists a sensored arc a that belongs to either path r or path r' and does not belong to the other and is equal to zero otherwise. $y''_{rr'}$ is equal to one if there is a pair of arcs a and a' such that both of them belong to both paths r and r' but their order of appearance differs between these paths; for example, in path r, arc a' follows arc a ($E_{raa'} = 1$), whereas in path r', arc a follows arc a' ($E_{r'aa'}$). Note that variables $y''_{rr'}$, $y'_{rr'}$ and $y_{rr'}$ are defined for all $r, r' \in R$ such that $r \ne r'$. $d(r, r', a)$ was defined earlier in the paper.

From constraints (9), (11) and (15), we find that $y_{rr'} = 1$ for all $r \ne r' \in R$. Thus, from constraint (8), we obtain $1 \le y'_{rr'} + y''_{rr'}$ for all $r \ne r' \in R$. In other words, the right-hand side of constraint (8) must equal one or more. Consider the network introduced in Example 1 and the case in which arcs 2 and 6 each have a sensor. In this case, constraints (6), (7) and (8) are as follows:

The right-hand side of constraint (6) is the number of sensored arcs that belong either to path r or to path r' and do not belong to the other. For example, paths 1 and 3 each contain two sensored arcs (i.e., arcs 2 and 6). Arc 2 belongs to path 1 and does not belong to path 3, whereas arc 6 belongs to path 3 and does not belong to path 1, and both arcs have a sensor. The right-hand side of constraint (7) is the number of pairs of sensored arcs that belong to both paths r and r' but appear in different orders in these two paths. For example, the sensored arcs 2 and 6 both belong to paths 2 and 4, and the orders in which they appear in paths 2 and 4 are different. Thus, there is one pair of sensored arcs that belongs to both paths 2 and 4 and appears in a different order in each of these paths. In this example, there is only one pair of paths that have at least two arcs in common. Therefore, there can be only one path pair for which the right-hand side of constraint (7) is greater than zero. For $r, r' \in R$, the right-hand side of constraint (8) is equal to at least one if and only if

Table 2
Constraints (6)-(8) for the network considered in Example 1 when arcs 2 and 6 each have a sensor.

Path pair $\{\boldsymbol{r,r'}\}$	Constraint (6)	Constraint (7)	Constraint (8)
{1,2}	$y'_{1,2} \le 1$	$y''_{1,2} \le 0$	$y_{1,2} \le 1+0$
{1,3}	$y'_{1,3} \le 2$	$y''_{1,3} \le 0$	$y_{1,3} \le 1+0$
{1,4}	$y'_{1,4} \le 1$	$y''_{1,4} \le 0$	$y_{1,4} \le 1+0$
{2,3}	$y'_{2,3} \le 1$	$y''_{2,3} \le 0$	$y_{2,3} \le 1+0$
{2,4}	$y'_{2,4} \le 0$	$y''_{2,4} \le 1$	$y_{2,4} \le 0+1$
{3,4}	$y'_{3,4} \le 1$	$y''_{3,4} \le 0$	$y_{3,4} \le 1+0$

the right-hand side of at least one of the constraints (6) and (7) is also equal to at least one.[5] As seen in the last column of Table 2, for all pairs of paths, the right-hand side of constraint (8) is equal to at least one. Constraint (7) considers the order of the sensored arcs. The role of this constraint and the sensor order can be seen from constraint (8) for the path pair {2, 4}. In this case, the right-hand side of constraint (6) is equal to zero, but because the right-hand side of constraint (7) is equal to one, the right-hand side of constraint (8) is also equal to one. The constraints of model L_1 are discussed further in the proof of Theorem 2.

Model L_1 is a location model that attempts to find a vector of sensor locations. As mentioned before, we have concluded that for every sensor location vector J, the set of path flows that satisfy condition (2) is non-empty; thus, the constraints of model L_1 are designed to ensure that this non-empty set is single-membered. In the following theorem, the necessity and sufficiency of the constraints of model L_1 will be claimed and proven correct.

Theorem 2. *Assume that J is the sensor location vector and that the order in which the vehicles are detected by the sensors is available. The constraints of model L_1 are sufficient and necessary conditions for the uniqueness of the path flows that satisfy condition (2).*

Proof of sufficiency. Suppose that the constraints of model L_1 are true; we wish to show that the path flows can then be uniquely defined.

For each pair of paths r and r', the right-hand side of constraint (6) is equal to one or more if there is at least one sensored arc for which $d(r, r', a) = 1$. As mentioned before, $d(r, r', a)$ is equal to one when arc a belongs either to path r or to path r' and does not belong to the other and is equal to zero otherwise. Thus, from constraint (6), we find that $y'_{rr'}$ can be greater than zero only when there exists a sensored arc a for which $d(r, r', a) = 1$. Considering the requirement $0 \le y'_{rr'} \le 1$ from constraint (15), if such an arc does not exist, we have $y'_{rr'} = 0$, and if such an arc exists, we have $y'_{rr'} \le 1$. Moreover, in the entire model, only constraints (6) and (15) impose an upper bound on $y'_{rr'}$.

Now, a property of constraints (12)–(15) will be claimed and proven so that it can be used in the remainder of the proof; according to these constraints, $J'_{aa'}$ has the following property:

$$J'_{aa'} = \begin{cases} 0 \text{ if } J_a = 0 \textit{ or } J_{a'} = 0 \\ 1 \text{ if } J_a = J_{a'} = 1 \end{cases}$$

All four possible cases are evaluated here: If $J_a = 1$ and $J_{a'} = 1$, then from constraints (12) and (13), we have $J'_{aa'} \le 1$, and from constraint (14), we have $J'_{aa'} \ge 1$; thus, we have $J'_{aa'} = 1$. If $J_a = 1$ and $J_{a'} = 0$ or if $J_a = 0$ and $J_{a'} = 1$, then from constraints (12) and (13), we have $J'_{aa'} \le 0$, and from constraint (14), we have $J'_{aa'} \ge 0$; thus, we have $J'_{aa'} = 0$. If $J_a = 0$ and $J_{a'} = 0$, then from constraints (12) and (13), we have $J'_{aa'} \le 0$, and from constraint (14), we have $J'_{aa'} \ge -1$; thus, because we know from constraint (15) that $J'_{aa'} \ge 0$, we have $J'_{aa'} = 0$.

The right-hand side of constraint (7) is a sum over all distinct arc pairs. For each arc pair a and a', the sum $(E_{raa'}E_{r'a'a} + E_{ra'a}E_{r'aa'})$ is equal to one if the following three constraints hold simultaneously:

- Both arcs a and a' have a sensor
- Both arcs a and a' belong to both paths r and r'
- These two arcs appear in these two paths in different orders

This happens in one of the following two cases:

- Both arcs a and a' have a sensor, arc a' follows arc a in path r ($E_{raa'} = 1$), and arc a follows arc a' in path r' ($E_{r'a'a} = 1$)
- Both arcs a and a' have a sensor, arc a follows arc a' in path r ($E_{ra'a} = 1$), and arc a' follows arc a in path r' ($E_{r'aa'} = 1$)

In the first case, the first term in the parentheses will be equal to one and the second will be zero. In the second case, the first term will be zero and the second will be equal to one. In either case, the value of the quantity in the parentheses will be one. Now, if $J'_{aa'} = 1$ (which means that $J_a = 1$ and $J_{a'} = 1$), there will be two sensored arcs that appear in both paths

[5] This constraint will be discussed more formally in the proof of Theorem 2.

r and r' but in different orders. In this case, the entire sum over all distinct arc pairs will be greater than or equal to one; therefore, constraint (7) together with constraint (15), from which we have $y''_{rr'} \le 1$, will result in $y''_{rr'} \le 1$. By contrast, if no such pair of arcs exists for paths r and r', then the right-hand side will be equal to zero. In this case, we will have $y''_{rr'} \le 0$, which, considering the requirement of $y''_{rr'} \ge 0$ from constraint (15), will result in $y''_{rr'} = 0$.

From constraints (9), (11) and (15), we find that $y_{rr'} = 1$ for all $r \ne r' \in R$. Thus, from constraint (8), we obtain $1 \le y'_{rr'} + y''_{rr'}$ for all $r \ne r' \in R$. As mentioned earlier, the upper bound on $y''_{rr'}$ is equal to one if there exists a pair of sensored arcs that appear in different orders in paths r and r' and is zero otherwise. Additionally, we note that the upper bound on $y'_{rr'}$ is equal to one if there exists a sensored arc a for which $d(r, r', a) = 1$ and is equal to zero otherwise. Thus, for each pair of paths $r \ne r' \in R$, the upper bound on at least one of $y'_{rr'}$ and $y''_{rr'}$ must be one. This is the case because if both upper bounds on $y'_{rr'}$ and $y''_{rr'}$ are zero, then both $y'_{rr'}$ and $y''_{rr'}$ must be zero, which contradicts the constraint $1 \le y'_{rr'} + y''_{rr'}$. Thus, for each pair of paths, one of the following clauses will be true:

- *First clause*: A sensored arc exists that belongs either to path r or to path r' and does not belong to the other.
- *Second clause*: A pair of sensored arcs exists in which both arcs belong to both paths r and r' and appear in different orders in these paths.

Paths r and r' are assigned to different combinations of sensors only when at least one of the above clauses is true. When the first clause is true, the sets of sensored arcs for paths r and r' are different and thus they correspond to different combinations. When the first clause is not true, the sets of sensored arcs for paths r and r' are the same; therefore, they correspond to different combinations only when the order of appearance of these sensored arcs differs between paths r and r'. This, in turn, is true only when there exists a pair of sensored arcs that belong to both paths and appear in different orders. Thus, in any arbitrary feasible solution for model L_1, any two paths r and r' correspond to different combinations of sensors.

From constraints (11) and (10), we find that $1 \le \sum_{a=1}^{n} \delta_{ar} J_a$, which means that each path has at least one sensored arc. Thus, if the constraints of model L_1 are satisfied, then all paths have at least one sensor and any two paths correspond to different combinations of sensors; therefore, the path flows can be uniquely determined.

Proof of necessity. The sensor location vector J is given. Suppose that the combination of sensors corresponding to each path is different from that corresponding to each other path and that each path has at least one sensor; now, we wish to prove that there exists some profile of x_r, $J'_{aa'}$, $y'_{rr'}$, $y''_{rr'}$ and $y_{rr'}$ for this J such that the constraints of model L_1 are true.

Consider two arbitrary paths r and r' that correspond to combinations $c \in C$ and $c' \in C$, respectively. We know that $c \ne c'$; thus, at least one of the first and second clauses stated above is true.

We claim that the following values for variables x_r, $J'_{aa'}$, $y'_{rr'}$, $y''_{rr'}$ and $y_{rr'}$ for the given J satisfy all constraints of model L_1:

$$y''_{rr'} = \begin{cases} 1 & \text{if the first clause is true for paths } r \text{ and } r' \\ 0 & \text{otherwise} \end{cases} \qquad \forall r, r' \in R | r \ne r'$$

$$y'_{rr'} = \begin{cases} 1 & \text{if the second clause is true for paths } r \text{ and } r' \\ 0 & \text{otherwise} \end{cases} \qquad \forall r, r' \in R | r \ne r'$$

$$y_{rr'} = \begin{cases} 1 & \text{if at least one of } y'_{rr'} \text{ and } y''_{rr'} \text{ is equal to one} \\ 0 & \text{otherwise} \end{cases} \qquad \forall r, r' \in R | r \ne r'$$

$$x_r = 1 \qquad \forall r \in R$$

$$J'_{aa'} = \begin{cases} 1 & \text{if } J_a = J_{a'} = 1 \\ 0 & \text{otherwise} \end{cases} \qquad \forall a, a' \in R | a \ne a'$$

The proof of this claim is trivial and is omitted here.

Lemma 1. *A solution to model L_1 always exists.*

Proof. Here, we claim and prove that the solution

$$\begin{cases} J_a = 1 & \forall a \in A \\ J'_{aa'} = 1 & \forall a \ne a' \in A \\ y_{rr'} = y'_{rr'} = 1 & \forall r \ne r' \in R \\ y''_{rr'} = 0 & \forall r \ne r' \in R \\ x_r = 1 & \forall r \in R \end{cases}$$

is always a feasible solution for model L_1, independent of the topological characteristics of the network and the set of routes considered.

If $J_a = 1$ for all $a \in A$, then for all $r \neq r' \in R$, the right-hand side of constraint (6) is equal to the number of arcs that belong to only one of these two paths. For every pair of paths, there exists at least one arc that belongs to only one of them. Thus, for all $r \neq r' \in R$, the right-hand side of constraint (6) is equal to at least one. Constraint (8) is obviously true. Therefore, in the proposed solution, constraints (6) and (8) are true for all $r \neq r' \in R$. For all path pairs, the right-hand side of constraint (7) is always nonnegative and the value of $y_{rr'}$ is zero. Therefore, constraint (7) is true for all $r \neq r' \in R$. Constraints (9) and (11) are obviously true. Because each path contains at least one arc, constraint (10) is also always true. Finally, in the proposed solution, constraints (12)–(16) will always be true. Thus, a solution for model L_1 always exists, independent of the topological characteristics of the network and the set of routes considered.

Lemma 1 is true because when vehicle identification sensors are used and every arc has a sensor, each set of paths can be uniquely observed. This is one of the primary advantages of these sensors compared with counting sensors.

3.1.3. Sensor location problem for optimal path flow observability considering a budget constraint

Usually, in practical examples, full observability of all path flows is not feasible because of limited resources. Thus, it is necessary to consider a budget constraint in the location model and attempt to maximize the coverage provided by the sensors. In this section, model L_2 is presented; this is a modified version of model L_1 that includes a budget constraint and maximizes the number of path flows that can be uniquely observed.

$$L_2: \ Maximize \sum_{r\in R} x_r$$

$$y'_{rr'} \le \sum_{a\in A} J_a d\left(r, r', a\right) \quad \forall r, r' \in R|\ r \neq r' \tag{17}$$

$$y''_{rr'} \le \sum_{a\in A} \sum_{\substack{a' \in A \\ a' \neq a}} J'_{aa'} (E_{raa'}E_{r'a'a} + E_{ra'a}E_{r'aa'}) \quad \forall r, r' \in R|\ r \neq r' \tag{18}$$

$$y_{rr'} \le y'_{rr'} + y''_{rr'} \quad \forall r, r' \in R|\ r \neq r' \tag{19}$$

$$x_r \le y_{rr'} \quad \forall r, r' \in R|\ r \neq r' \tag{20}$$

$$x_r \le \sum_{a\in A} \delta_{ar} J_a \quad \forall r \in R \tag{21}$$

$$J'_{aa'} \le J_a \quad \forall a, a' \in A|\ a \neq a' \tag{22}$$

$$J'_{aa'} \le J_{a'} \quad \forall a, a' \in A|\ a \neq a' \tag{23}$$

$$J'_{aa'} \ge J_a + J_{a'} - 1 \quad \forall a, a' \in A|\ a \neq a' \tag{24}$$

$$\sum_{a\in A} J_a \le B \tag{25}$$

$$0 \le x_r,\ J'_{aa'}, y'_{rr'}, y''_{rr'}, y_{rr'} \le 1 \quad \forall r, r' \in R, \quad \forall a, a' \in A \tag{26}$$

$$J_a = 0, 1 \quad \forall a \in A \tag{27}$$

B is the available sensor budget (i.e., the number of available sensors). The remaining variables and parameters are similar to those of model L_1. There are three differences between models L_1 and L_2: The first difference is that constraint (11), which guarantees the full observability of all paths, is omitted. The second difference is the target function, which serves to maximize the number of fully observable paths. The third difference is that constraint (25), which is the budget constraint, has been added. With the exception of constraint (25), all constraints in model L_2 are also included in model L_1 and were explained in the previous section. Below, two lemmas regarding model L_2 are presented.

Lemma 2. *In model L_2, x_r is equal to one if the flow on path r can be uniquely defined and is zero otherwise.*

Lemma 3. *A solution to model L_2 always exists.*

Based on the facts established in the proof of Theorem 2, the proof of Lemma 2 is trivial and is omitted here. For the proof of Lemma 3, we simply present the following solution, which is always a feasible solution for model L_2:

$$\begin{cases} J_a = 0 \quad \forall a \in A \\ J'_{aa'} = 0 \quad \forall a \neq a' \in A \\ y_{rr'} = y'_{rr'} = 0 \quad \forall r \neq r' \in R \\ y''_{rr'} = 0 \quad \forall r \neq r' \in R \\ x_r = 0 \quad \forall r \in R \end{cases}$$

It is obvious that in the target function, we can choose to maximize the net flow instead of the number of uniquely observed paths; this can easily be done by multiplying x_r by f_r^0, an out-of-date flow for path r, for all $r \in R$.

3.2. Sensor location problem for the uniqueness of the OD vector

For OD vector estimation, observing all path flows is sufficient but not always necessary. Based on model C, model D is proposed below for determining the locations of vehicle identification sensors. In model D, the constraints are sufficient but not necessary to ensure the uniqueness of the OD vector.

$$D: \ \textit{Minimize } n_s = \sum_{a \in A} J_a$$

$$J_a \in \{0, 1\} \quad \forall a \in A \tag{28}$$

$$\sum_{a \in A} J_a d(r_1, r_2, a) \geq 1 \quad \forall i, i' \in I | i \neq i', \quad \forall r_1 \in R_i, \quad \forall r_2 \in R_{i'} \tag{29}$$

$$\sum_{a \in A} J_a \delta_{ar} \geq 1 \quad \forall r \in R \tag{30}$$

The only difference between model D and model C is the pairs of paths for which constraint (29) must be true. In model C, constraint (4) must hold for all possible pairs of paths, whereas in model D, constraint (29) must hold only for those pairs of paths that belong to different OD pairs. Model D is a location model, and for its solution, which is a sensor location vector, the set of path flows and OD vectors that satisfy condition (2) is non-empty. A theorem regarding model D is presented below.

Theorem 3. *Assume that J is the sensor location vector and that the order in which the vehicles are detected by the sensors is available. The constraints of model D are sufficient but not necessary conditions for the uniqueness of the OD vector that satisfies condition (2).*

Proof of sufficiency. It has previously been concluded that the set of path flows and OD vectors that satisfy condition (2) is always non-empty. To prove this theorem, we must show that exactly one OD vector satisfies condition (2) when constraints (29) and (30) are true. We define

$$C_i := \textit{the set of all combinations that correspond to at least one of the paths}$$

$$\textit{belonging to OD pair } i$$

From (1), we know that each path corresponds to at most one combination. From constraint (29), we know that for every two paths with different OD pairs, there exists at least one sensored arc that belongs to one of them and does not belong to the other. Thus, these two paths correspond to different combinations. Constraint (29) allows only paths with the same OD pair to correspond to the same combination. Thus, we have

$$C_i \cap C_{i'} = \emptyset \quad \forall i, i' \in I \ | i \neq i' \tag{31}$$

To understand the reason for relation (31), suppose that C_i and $C_{i'}$ have an element in common.

Relation (31) holds because if C_i and $C_{i'}$ have an element in common, then a combination must simultaneously correspond to a path belonging to OD pair i and a path belonging to OD pair i'. This contradicts constraint (29).

According to constraint (30), all paths in the network have at least one sensor and correspond to a non-empty combination. Thus, an empty combination does not belong to $\cup_{i \in I} C_i$ Each of the paths belonging to OD pair i corresponds to one of the elements of C_i, and $R_i \subseteq (\cup_{j \in C_i} R'_j)$. From (31), we know that none of the paths that belong to any OD pair other than i corresponds to any of the elements of C_i. Thus, all paths that correspond to the elements of C_i belong to OD pair i. Therefore, $\cup_{j \in C_i} R'_j \subseteq R_i$. Consequently,

$$\left(\underset{j \in C_i}{\cup} R'_j \right) = R_i \tag{32}$$

Here, let us recall condition (2):

$$\sum_{r\in R'_j} f_r = b_j \quad \forall j \in C | j \neq \emptyset \tag{2a}$$

$$\sum_{r\in R_i} f_r = Q_i \quad \forall i \in I, \quad \forall r \in R \tag{2b}$$

In constraint (2b), if we replace R_i with its equivalent expression from (32), we obtain

$$Q_i = \sum_{r\in R_i} f_r = \sum_{r\in (\underset{j\in C_i}{\cup} R'_j)} f_r = \sum_{j\in C_i}\sum_{r\in R'_j} f_r = \sum_{j\in C_i} b_j \tag{33}$$

The first equation is obtained from (2b). The second equation is obtained from (32). The third equation is obtained from (1). The fourth equation is obtained from (2a) and the fact that none of the combinations $j \in C_i$ is an empty set, which was established earlier in the proof. Thus, the only OD vector that satisfies condition (2) is the OD vector obtained in (33).

To show that the constraints of model D are not necessary, we can use Example 1. In this example, the solution to model D is {2, 3, 6}, but we know that {2, 6} will also result in unique path flows and a unique OD vector. Notably, the solution for which $J_a = 1$ for all $a \in A$ is always a feasible solution for model D. Thus, model D always has a solution.

3.2.1. Necessary and sufficient conditions for the uniqueness of the OD vector

According to Theorem 3, the constraints of model D are sufficient but not necessary conditions for the uniqueness of the OD vector. Based on the underlying concepts of model L_1, model L_3 is proposed below to ensure the necessary and sufficient conditions for the uniqueness of the OD vector. This property will be claimed formally in Theorem 4.

$$L_3: \ \textit{Minimize } n_s = \sum_{a\in A} J_a$$

$$y'_{rr'} \le \sum_{a\in A} J_a d\left(r, r', a\right) \quad \forall i, i' \in I, \quad \forall r \in R_i, \quad \forall r' \in R_{i'} |\ i \neq i' \tag{34}$$

$$y''_{rr'} \le \sum_{a\in A} \sum_{\substack{a' \in A \\ a' \neq a}} J'_{aa'}(E_{raa'}E_{r'a'a} + E_{ra'a}E_{r'aa'}) \quad \forall i, i' \in I, \quad \forall r \in R_i, \quad \forall r' \in R_{i'} |\ i \neq i' \tag{35}$$

$$y_{rr'} \le y'_{rr'} + y''_{rr'} \quad \forall i, i' \in I, \quad \forall r \in R_i, \quad \forall r' \in R_{i'} |\ i \neq i' \tag{36}$$

$$x_r \le y_{rr'} \quad \forall i, i' \in I, \quad \forall r \in R_i, \quad \forall r' \in R_{i'} |\ i \neq i' \tag{37}$$

$$x_r \le \sum_{a\in A} \delta_{ar} J_a \quad \forall r \in R \tag{38}$$

$$x_r = 1 \quad \forall r \in R \tag{39}$$

$$J'_{aa'} \le J_a \quad \forall a, a' \in A |\ a \neq a' \tag{40}$$

$$J'_{aa'} \le J_{a'} \quad \forall a, a' \in A |\ a \neq a' \tag{41}$$

$$J'_{aa'} \ge J_a + J_{a'} - 1 \quad \forall a, a' \in A |\ a \neq a' \tag{42}$$

$$0 \le x_r,\ J'_{aa'}, y'_{rr'}, y''_{rr'}, y_{rr'} \le 1 \quad \forall i, i' \in I, \quad \forall r \in R_i, \quad \forall r' \in R_{i'} |\ i \neq i', \quad \forall a, a' \in A \tag{43}$$

$$J_a = 0, 1 \quad \forall a \in A \tag{44}$$

The only difference between models L_1 and L_3 is the pairs of paths for which constraints (34)–(37) must be true. In model L_1, these constraints must hold for every possible pair of paths. However, in model L_3, these constraints must hold only for pairs of paths that do not belong to the same OD pair. Constraints (34)–(37) guarantee that any two paths with different OD pairs will correspond to different combinations of sensors. Consequently, variables $y''_{rr'}$, $y'_{rr'}$ and $y_{rr'}$ are defined for all $r \in R_i$ and $r' \in R_{i'}$ such that $i \neq i' \in I$. Thus, in model L_3, the following statements hold for all $r \in R_i$ and $r' \in R_{i'}$ such that $i \neq i' \in I$:

- $y_{rr'}$ is equal to one if paths r and r' correspond to different combinations and is zero otherwise
- $y'_{rr'}$ is equal to one if there exists a sensored arc that belongs either to path r or to path r' and does not belong to the other
- $y''_{rr'}$ is equal to one if there exist two sensored arcs that appear in both paths r and r' in different orders

Consider Example 1, in which there are two OD pairs and two paths exists for each OD pair. Paths 1 and 2 belong to the OD pair 1–3, and paths 3 and 4 belong to the OD pair 2–4. In this case, variables $y''_{rr'}$, $y'_{rr'}$ and $y_{rr'}$ are defined for the following path pairs: {1,3}, {1,4}, {2,3} and {2,4}. These are all possible pairs of paths whose elements belong to different OD pairs. Constraints (34)–(37) must hold only for these four path pairs.

A theorem regarding this model is presented below:

Theorem 4. *Assume that J is the sensor location vector and that the order in which the vehicles are detected by the sensors is available. The constraints of model L_3 are necessary and sufficient conditions for the uniqueness of the OD vector that satisfies condition (2).*

The proof of this theorem is similar to that of Theorem 2 and is omitted here. A lemma similar to Lemma 1 is also presented here.

Lemma 4. *A solution to model L_3 always exists.*

The proof of this lemma is similar to that of Lemma 1 and is omitted here.

3.2.2. Sensor location problem for optimal OD flow observability considering a budget constraint

In this section, model L_4 is presented; this is a modified version of model L_3 that includes a budget constraint and maximizes the number of OD pair flows that can be uniquely observed.

$$L_4: \ \text{Maximize} \sum_{i\in I} z_i$$

$$y'_{rr'} \le \sum_{a\in A} J_a d(r, r', a) \quad \forall i, i' \in I, \quad \forall r \in R_i, \quad \forall r' \in R_{i'} |\ i \neq i' \tag{45}$$

$$y''_{rr'} \le \sum_{a\in A} \sum_{\substack{a' \in A \\ a' \neq a}} J'_{aa'} (E_{raa'}E_{r'a'a} + E_{ra'a}E_{r'aa'}) \quad \forall i, i' \in I, \quad \forall r \in R_i, \quad \forall r' \in R_{i'} |\ i \neq i' \tag{46}$$

$$y_{rr'} \le y'_{rr'} + y''_{rr'} \quad \forall i, i' \in I, \quad \forall r \in R_i, \quad \forall r' \in R_{i'} |\ i \neq i' \tag{47}$$

$$x_r \le y_{rr'} \quad \forall i, i' \in I, \quad \forall r \in R_i, \quad \forall r' \in R_{i'} |\ i \neq i' \tag{48}$$

$$x_r \le \sum_{a\in A} \delta_{ar} J_a \quad \forall r \in R \tag{49}$$

$$z_i \le x_r \quad \forall i \in I, \quad \forall r \in R_i \tag{50}$$

$$J'_{aa'} \le J_a \quad \forall a, a' \in A |\ a \neq a' \tag{51}$$

$$J'_{aa'} \le J_{a'} \quad \forall a, a' \in A |\ a \neq a' \tag{52}$$

$$J'_{aa'} \ge J_a + J_{a'} - 1 \quad \forall a, a' \in A |\ a \neq a' \tag{53}$$

$$\sum_{a\in A} J_a \le B \tag{54}$$

$$0 \le x_r, \ J'_{aa'}, y'_{rr'}, y''_{rr'}, y_{rr'}, z_i \le 1 \quad \forall r, r' \in R, \quad \forall a, a' \in A \tag{55}$$

$$J_a = 0, 1 \quad \forall a \in A \tag{56}$$

Suppose that path r corresponds to combination $j \in C$ and that its OD pair is i. In the above model, x_r is equal to one if there is at least one sensor on path r and $R'_j \subseteq R_i$. Model L_4 is designed such that z_i is equal to one if each path in R_i has a sensor and is not associated with the same combination as any path not in R_i and is zero otherwise. Thus, z_i is equal to one if and only if the flow for OD pair i can be uniquely defined. Therefore, the target function of model L_4 maximizes the number of OD flows that can be uniquely defined. It is possible to multiply each z_i by the importance of the corresponding OD pair i in the target function. A lemma regarding this model is presented here:

Lemma 5. *A solution to model L_4 always exists.*

The proof of this lemma is similar to that of Lemma 3 and is omitted here.

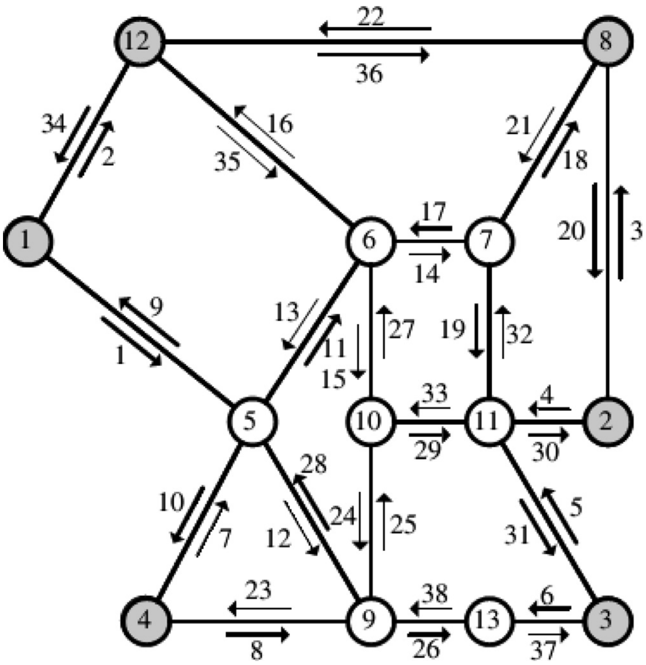

Fig. 2. Nguyen–Dupuis network (Mínguez et al., 2010).

4. Numerical results

This section presents several examples of the application and solution of models L_1, L_2, L_3 and L_4 to address various numerical aspects of these models. Specifically, the solutions to four examples, examples 2–5, are reported in this section. Example 2 uses the topology of the Nguyen–Dupuis network, with 38 arcs; examples 3 and 4 use the topology of the Sioux–Falls network, with 76 arcs; and example 5 uses the topology of the Mashhad network, with 2526 arcs. Examples 2 and 3 were solved for models L_1, L_2, L_3 and L_4. However, because of the high CPU time required, example 4 was solved only for models L_1 and L_3 and example 5 was solved only for model L_1. All solutions reported in this section were obtained using GAMS on a computer with a Core i5 CPU clocking at 2.53 GHz. The solver used by GAMS was CPLEX.

4.1. Example 2: Nguyen–Dupuis network

In this example, the Nguyen–Dupuis test network is considered. This network, which contains 38 arcs, is depicted in Fig. 2. We consider the same set of OD pairs and routes considered by Mínguez et al. (2010). Their example includes 18 OD pairs and 50 routes.

4.1.1. Model L_1

Model L_1 yields the minimum number of sensors required for full observability of the path flows. For this example, this model was solved for the Nguyen–Dupuis network, which contains 18 OD pairs and 50 routes. The CPU time required to solve this problem was less than 1 s, and the solution consists of 18 sensors:

$$\{1, 2, 3, 5, 8, 9, 11, 13, 18, 20, 21, 22, 23, 29, 31, 33, 34, 36\}$$

Thus, the 50 considered path flows in this network can be fully observed by 18 vehicle identification sensors.

4.1.2. Model L_2

Model L_2 answers the following question: how many of the considered paths can be fully observed with a given sensor budget? Model L_2 was solved for sensor budgets of 1–18. As we know from the results of model L_1, a budget of 18 sensors is sufficient for full observability of all paths. Fig. 3 shows the percentage of the considered paths that can be fully observed as a function of the budget. Fig. 4 shows the CPU times required to solve the model for different budget levels. The maximum CPU time, 40 s, was consumed for a budget of 11.

4.1.3. Model L_3

Model L_3 yields the minimum number of sensors required for full observability of the OD flows; an OD flow is fully observable when it can be uniquely defined based on the observations provided by the sensors. This example, which includes

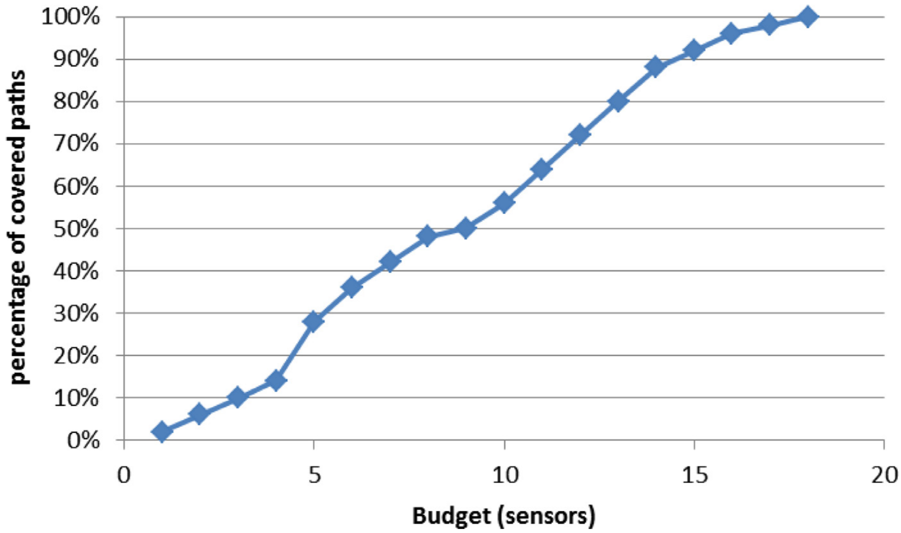

Fig. 3. Percentage of paths covered in example 2 as a function of the sensor budget.

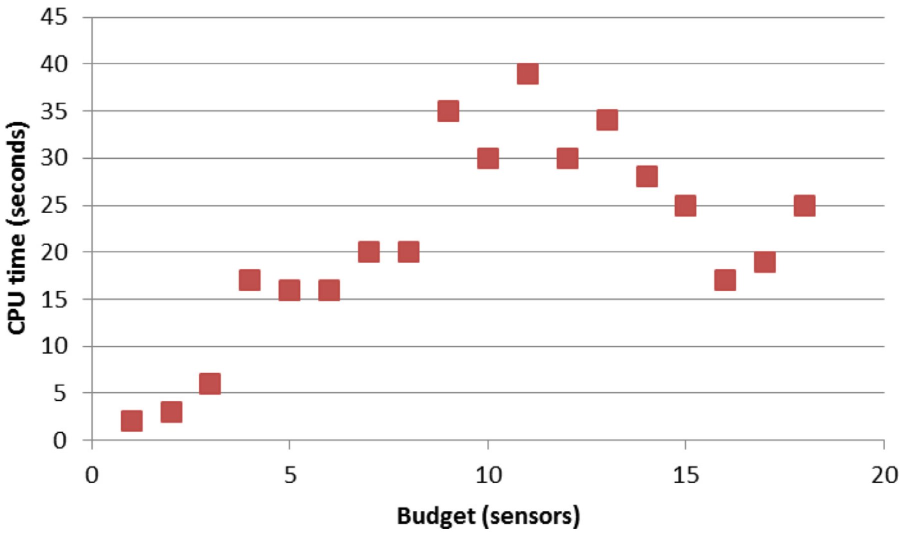

Fig. 4. CPU times required to solve model $\boldsymbol{L}_2$ for example 2 given different budget levels.

18 OD pairs, was solved in less than 1 s of CPU time; the solution is as follows:

$$\{1, 2, 3, 4, 8, 9, 16, 19, 20, 21, 22, 23, 30, 34, 35, 36\}$$

This solution consists of 16 sensors; thus, for full observability of the OD pairs in this example, 16 sensors are necessary and sufficient.

4.1.4. Model L_4

There are 18 OD pairs in this example; from the solution of model L_3, we know that for full observability of these OD flows, 16 sensors are necessary and sufficient. Model L_4 yields the maximum number of OD flows that are fully observable with a given sensor budget. Fig. 5 shows the number of OD pairs that can be fully observed as a function of the budget. Fig. 6 shows the CPU times required to solve model L_4 for example 2 for different budget levels.

4.2. Example 3: Upper half of the Sioux–Falls network

In this example, the topology of the Sioux–Falls test network is considered. This network, which contains 76 arcs, is depicted in Fig. 7. Six OD pairs are considered: 1–17, 17–1, 3–18, 18–3, 12–2 and 2–12; these OD pairs and their corresponding paths are represented in Table 3. Ninety-two paths are considered for this network; in Table 4, the arc set for each path is

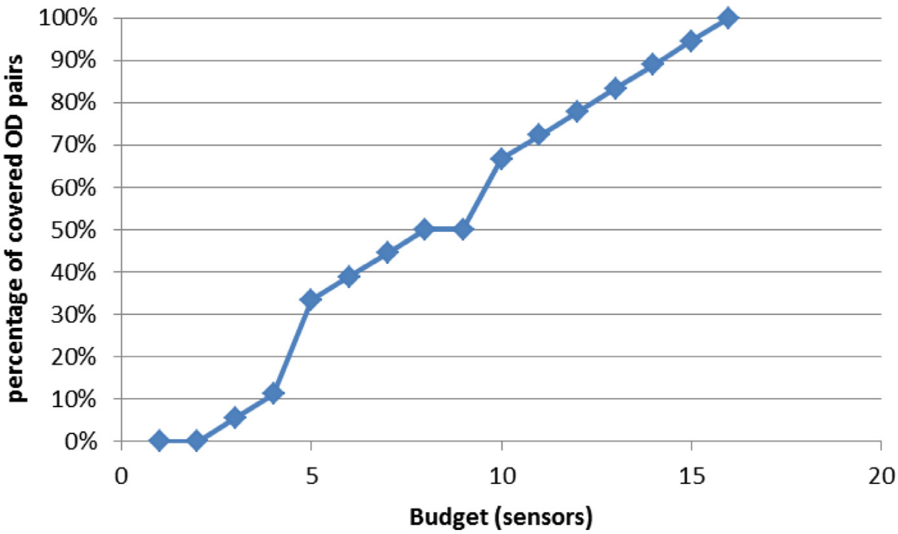

Fig. 5. Number of fully observable OD flows in example 2 as a function of the sensor budget.

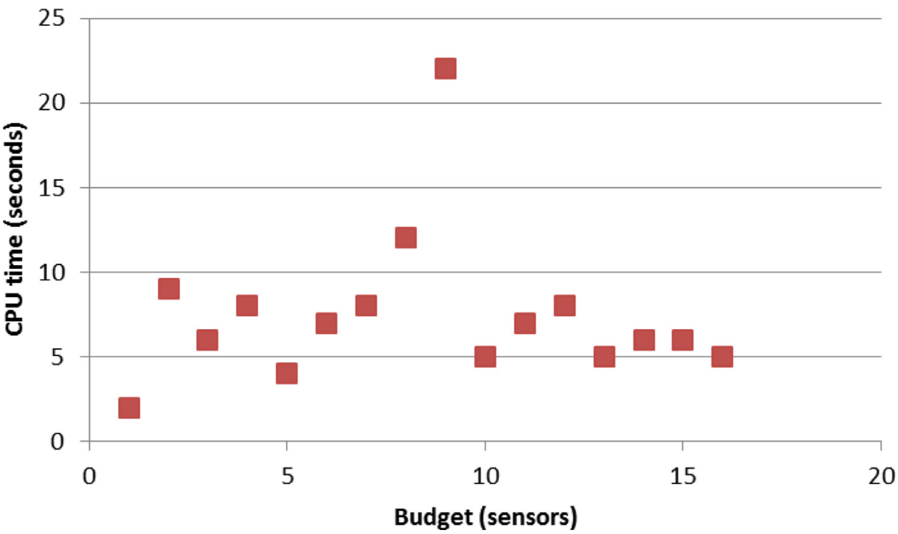

Fig. 6. CPU times required to solve model $\boldsymbol{L_4}$ for example 2 given different budget levels.

Table 3
Set of paths considered for each OD pair in example 3.

OD pair	1–17	17–1	3–18	18–3	12–2	2–12
Paths	1, 2, ..., 20	21, 22, ..., 40	41, 42, ..., 55	56, 57, ..., 70	71, 72, ..., 81	82, 83, ..., 92

presented. For each OD pair, all reasonable simple paths are considered; by reasonable paths, we mean paths that are not considerably longer than the shortest path between the origin and the destination. In this example, because the OD pairs are all located in the upper half of the network, there are only 52 arcs that are included in at least one path. Thus, although the Sioux–Falls network contains 76 arcs in total, the maximum practical number of sensor locations in this example is 52.

4.2.1. Model L_1

Model L_1 yields the minimum number of sensors required for full observability of the path flows. For this example, this model was solved considering all 92 paths represented in Table 4. The CPU time required to solve this problem was 15 min, and the solution consists of 18 sensors:

$$\{2, 4, 5, 6, 8, 13, 14, 22, 23, 27, 29, 30, 32, 33, 36, 47, 48, 51\}$$

Thus, the 92 path flows considered in this network can be fully observed by 18 vehicle identification sensors.

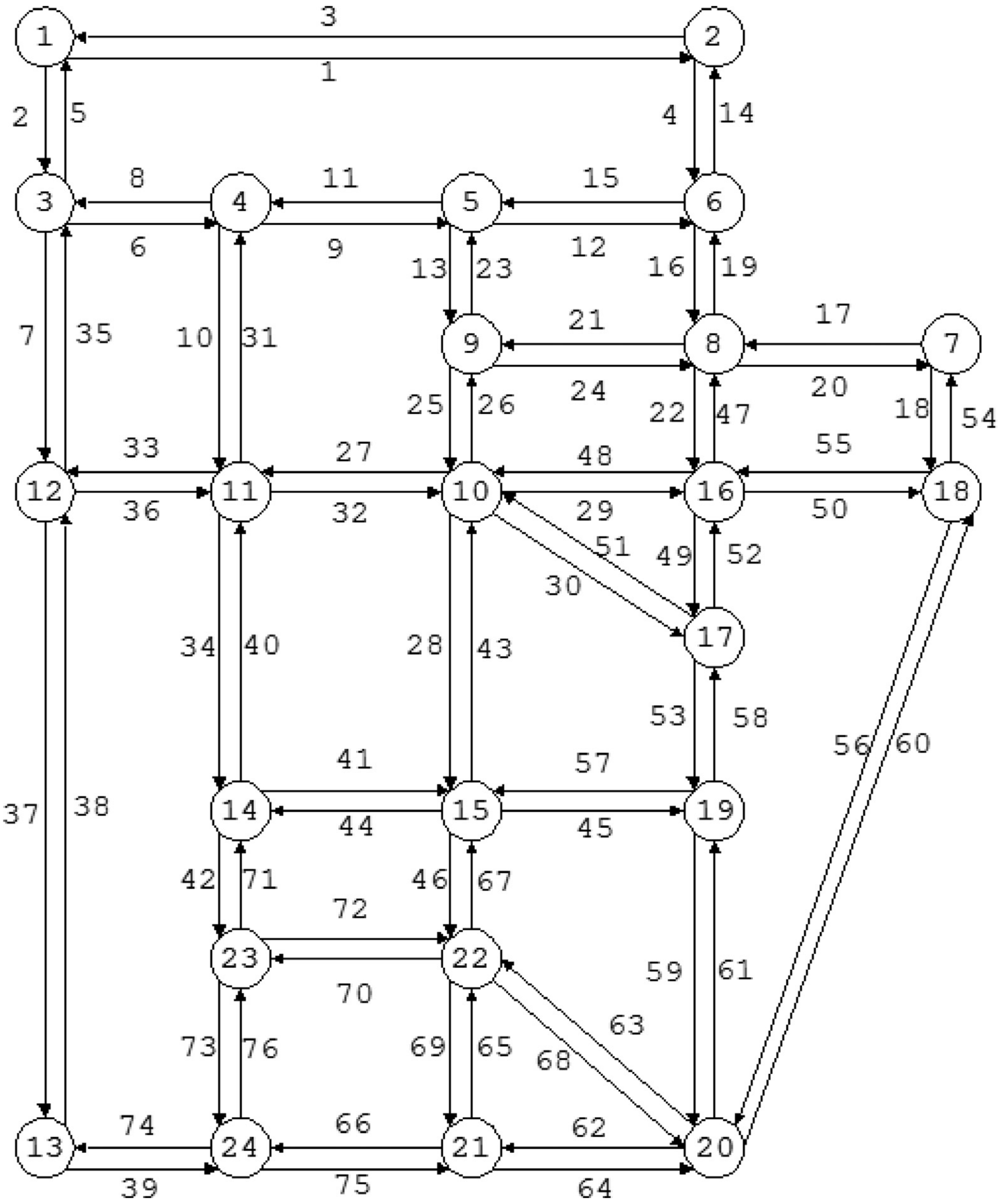

Fig. 7. Sioux–Falls network.

4.2.2. *Model* L_2

Model L_2 answers the following question: how many of the considered paths can be fully observed with a given sensor budget? Model L_2 was solved for sensor budgets of 1–18. As we know from the results of model L_1, a budget of 18 sensors is sufficient for full observability of all paths. Fig. 8 shows the percentage of the considered paths that can be fully observed as a function of the budget. Fig. 9 shows the CPU times required to solve the model for different budget levels. The maximum CPU time, 192 min, was consumed for a budget of 17.

4.2.3. *Model* L_3

Model L_3 yields the minimum number of sensors required for full observability of the OD flows. This example, which includes 6 OD pairs, was solved within 5 s of CPU time; the solution is as follows:

$$\{2, 4, 5, 6, 8, 14, 33, 36\}$$

This solution consists of eight sensors; thus, for full observability of the OD pairs in this example, eight sensors are necessary and sufficient. The CPU time required to solve this model is less than that for model L_1; the reason is that model

Table 4
Paths in the Sioux–Falls network considered in example 3.

Path	Links	Path	Links	Path	Links
1	2 7 36 32 30	32	51 26 24 19 14 3	63	55 49 53 57 44 40 33 35
2	2 7 36 31 9 13 25 30	33	51 29 47 19 14 3	64	55 47 21 23 11 8
3	2 7 36 34 41 43 30	34	51 29 50 54 17 19 14 3	65	54 17 21 23 11 8
4	2 6 9 13 25 30	35	51 27 31 9 12 14 3	66	55 47 19 15 11 8
5	2 6 10 32 30	36	51 29 47 21 23 12 14 3	67	54 17 19 15 11 8
6	2 6 9 12 16 21 25 30	37	52 47 21 23 12 14 3	68	55 47 19 14 3 2
7	2 6 9 12 16 22 48 30	38	52 50 54 17 19 14 3	69	54 17 19 14 3 2
8	2 6 9 12 16 22 49	39	53 57 44 40 33 35 5	70	55 48 26 24 19 14 3 2
9	2 6 9 13 24 22 48 30	40	53 57 44 40 31 8 5	71	35 5 1
10	2 6 9 13 24 20 18 55 48 30	41	7 36 32 29 50	72	35 6 9 12 14
11	1 4 15 13 25 30	42	6 10 32 29 50	73	35 6 9 13 24 19 14
12	1 4 16 21 25 30	43	6 9 13 25 29 50	74	36 31 9 12 14
13	1 4 16 22 48 30	44	7 36 34 41 43 29 50	75	36 31 9 13 24 19 14
14	1 4 16 20 18 55 48 30	45	7 36 32 30 52 50	76	36 32 26 23 12 14
15	1 4 15 11 10 32 30	46	6 10 32 30 52 50	77	36 32 26 24 19 14
16	1 4 15 13 24 22 48 30	47	6 9 13 25 30 52 50	78	36 32 29 47 19 14
17	1 4 15 13 24 22 49	48	7 36 34 41 45 58 52 50	79	36 32 30 52 47 19 14
18	1 4 16 20 18 55 49	49	6 9 13 24 22 50	80	36 32 29 50 54 17 19 14
19	2 7 36 34 41 45 58	50	6 9 13 24 20 18	81	36 32 29 47 21 23 12 14
20	2 6 10 34 41 45 58	51	6 9 12 16 22 50	82	3 2 7
21	51 27 33 35 5	52	6 9 12 16 20 18	83	4 5 11 8 7
22	51 26 23 11 10 33 35 5	53	5 1 4 16 22 50	84	4 16 21 23 11 8 7
23	51 28 44 40 33 35 5	54	5 1 4 16 20 18	85	4 15 11 10 33
24	51 26 23 11 8 5	55	5 1 4 16 21 25	86	4 16 21 23 11 10 33
25	51 27 31 8 5	56	55 48 27 33 35	87	4 15 13 25 27 33
26	51 26 24 19 15 11 8 5	57	55 48 27 31 8	88	4 16 21 25 27 33
27	51 29 47 19 15 11 8 5	58	55 48 26 23 11 8	89	4 16 22 48 27 33
28	52 47 19 15 11 8 5	59	55 48 28 44 40 33 35	90	4 16 22 49 51 27 33
29	51 29 47 21 23 11 8 5	60	55 49 51 27 33 35	91	4 16 20 18 55 48 27 33
30	51 29 50 54 17 21 23 11 8 5	61	55 49 51 27 31 8	92	4 15 13 24 22 48 27 33
31	51 26 23 12 14 3	62	1 26 23 11 8		

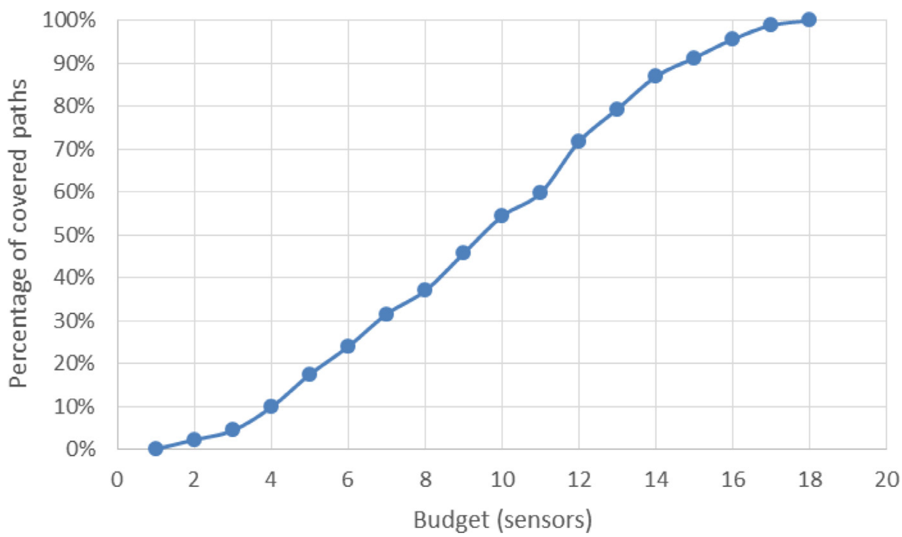

Fig. 8. Percentage of paths covered in example 3 as a function of the sensor budget.

L_3 has fewer constraints than does model L_1. In this example, model L_1 has 25,478 constraints, whereas model L_3 has 11,534 constraints.

4.2.4. Model L_4

There are six OD pairs in this example; from the solution of model L_3, we know that for full observability of these OD flows, 8 sensors are necessary and sufficient. Model L_4 yields the maximum number of OD flows that are fully observable with a given sensor budget. Fig. 10 shows the number of OD pairs that can be fully observed as a function of the budget. Fig. 11 shows the CPU times required to solve model L_4 for example 3 for different budget levels.

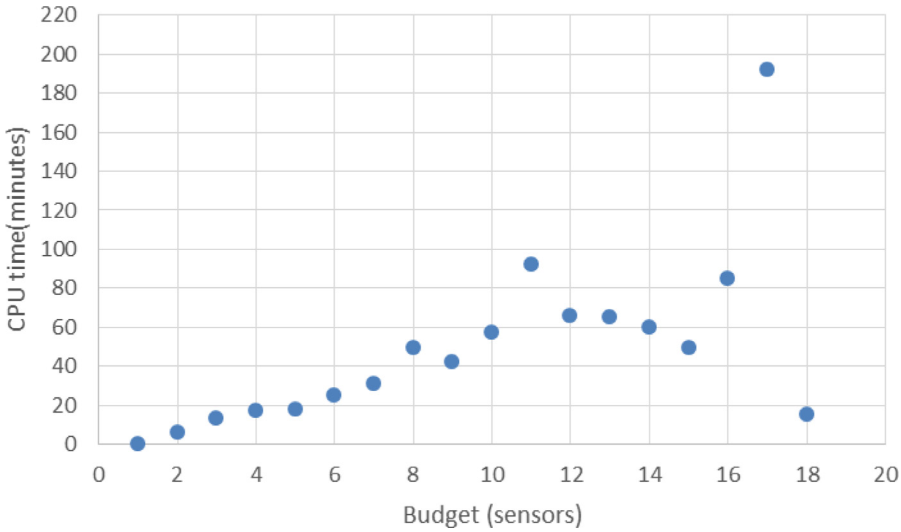

Fig. 9. CPU times used to solve model $\boldsymbol{L}_2$ for example 3 given different budget levels.

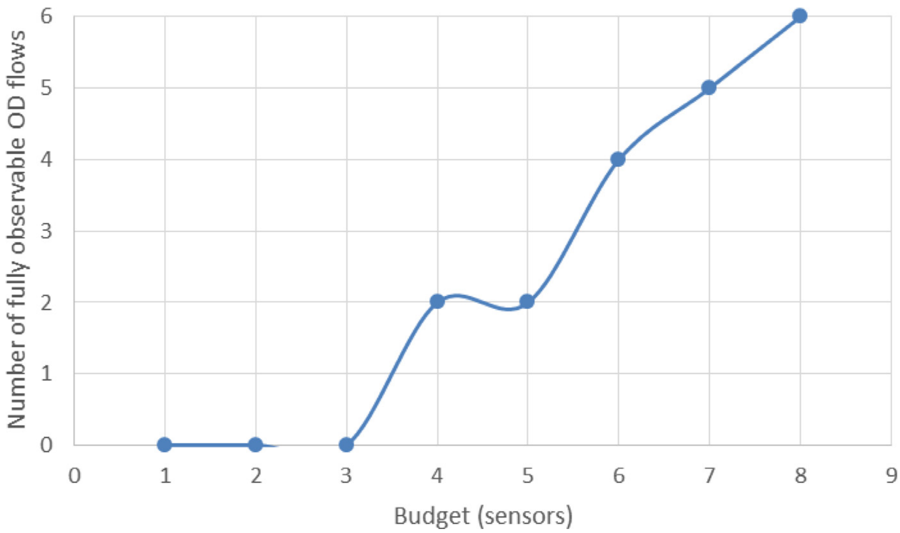

Fig. 10. Number of fully observable OD flows in example 3 as a function of the sensor budget.

4.3. Example 4: Complete Sioux–Falls network

In this example, the topology of the Sioux–Falls test network is again considered. This example consists of four cases. In these cases, the number of OD pairs remains fixed at 12, but the numbers of considered paths are 112, 156, 424 and 778 for cases 1–4, respectively. The 12 OD pairs are 1–20, 20–1, 3–18, 18–3, 12–7, 7–12, 13–8, 8–13, 24–6, 6–24, 21–2 and 2–21. In each case, all possible paths for each OD pair that are no longer than some predefined proportion of the length of the shortest path between that OD pair are considered. This predefined proportion takes values of 0.05, 0.1, 0.3 and 0.4 in cases 1–4, respectively. The results of solving models L_1 and L_3 for these four cases are reported in this section.

4.3.1. Model L_1

Model L_1 yields the minimum number of sensors required for full observability of the path flows. This model was solved for cases 1–3 of example 4. The fourth case could not be solved within 1 week, and therefore, no result for this case is reported here. The three obtained solutions are as follows:

$$\{7, 8, 9, 10, 13, 17, 20, 22, 23, 27, 29, 32, 35, 39, 46, 47, 48, 66, 67, 73, 74, 75, 76\}$$

$$\{7, 10, 12, 15, 16, 17, 19, 20, 22, 27, 28, 29, 31, 32, 35, 43, 46, 47, 48, 59, 61, 62, 64, 66, 71, 73, 75, 76\}$$

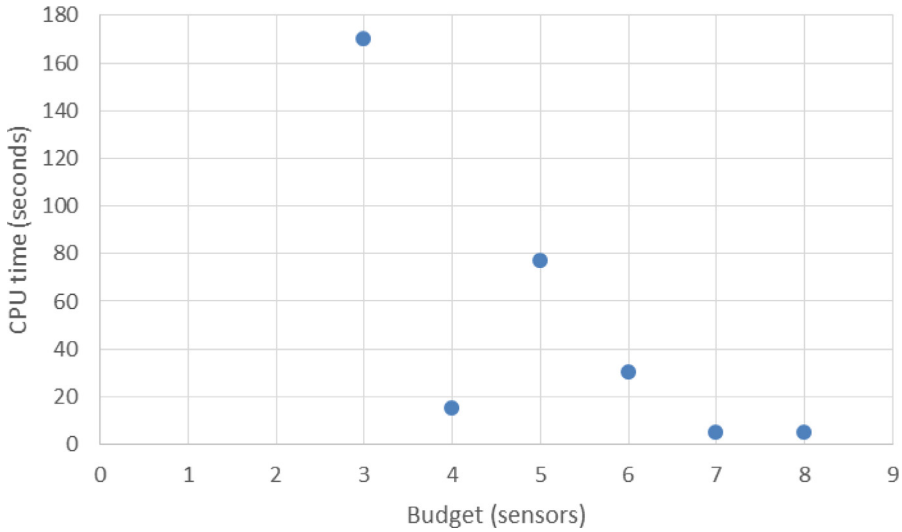

Fig. 11. CPU times required to solve model $\boldsymbol{L}_4$ for example 3 given different budget levels.

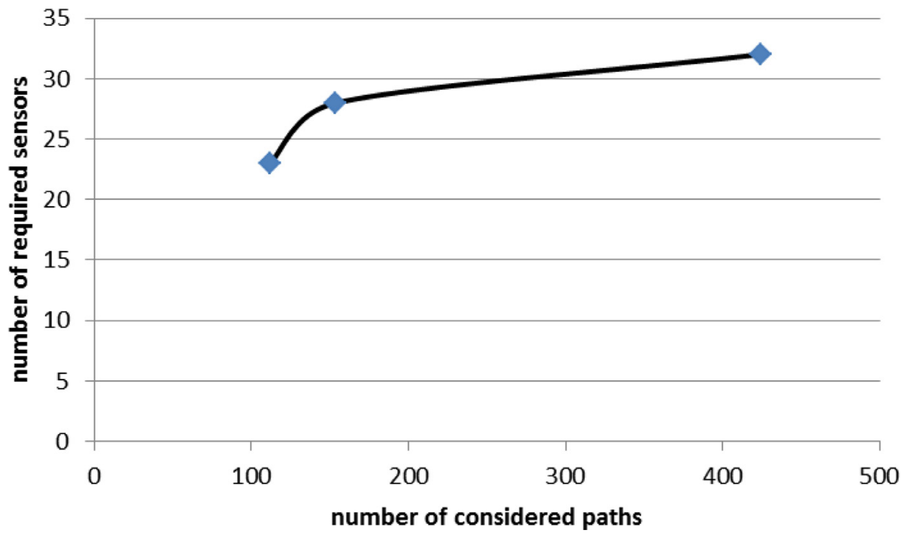

Fig. 12. Optimal solutions of model $\boldsymbol{L}_1$ for cases 1–3 of example 4.

$$\{2, 4, 5, 7, 10, 13, 14, 16, 19, 21, 22, 23, 24, 27, 29, 31, 32, 35, 37, 38, 42, 46, 47, 48, 53, 58, 62, 64, 65, 67, 69, 71\}$$

Fig. 12 shows the numbers of required sensors for these three cases. Fig. 13 shows the CPU times required to solve model L_1 for these cases. In Fig. 13, the CPU time required to solve the model for 112 paths is 5 min, whereas the time required for 154 paths is 104 min and that required for 424 paths is 3367 min, or 2.3 days. Thus, the time required to solve model L_1 grows extremely rapidly as the number of considered paths increases. Fig. 12 shows that the accompanying rate of increase in the required number of sensors is considerably milder.

In medium- and large-scale problems, the amounts of memory required to solve these models can be problematic. Specifically, the two three-dimensional matrices $d(r, r', a)$ and $E(r, a, a')$ can consume large amounts of memory in the case of large networks. These two matrices appear in all of our location models. The first matrix, $d(r, r', a)$, can be calculated in closed form based on the arc-route matrix δ: $d(r, r', a) = (\delta_{ar} + \delta_{ar'}) * (1 - \delta_{ar}\delta_{ar'})$; This relation has previously been presented by Mínguez et al. (2010). Thus, there is no need to pass matrix d as an input.

Matrix E is the other three-dimensional matrix. This matrix appears in the following constraint:

$$y''_{rr'} \leq \sum_{a \in A} \sum_{\substack{a' \in A \\ a' \neq a}} J'_{aa'} (E_{raa'} E_{r'a'a} + E_{ra'a} E_{r'aa'})$$

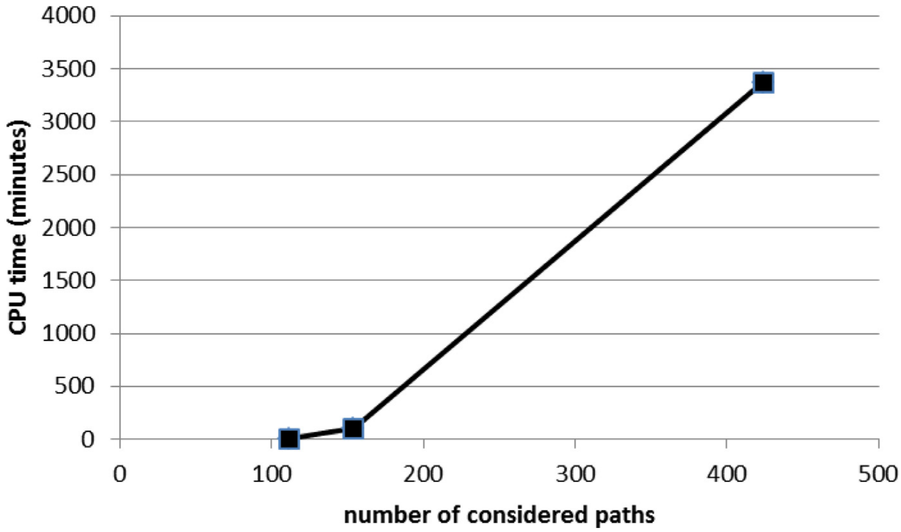

Fig. 13. CPU times required to solve model L_1 for cases 1–3 of example 4.

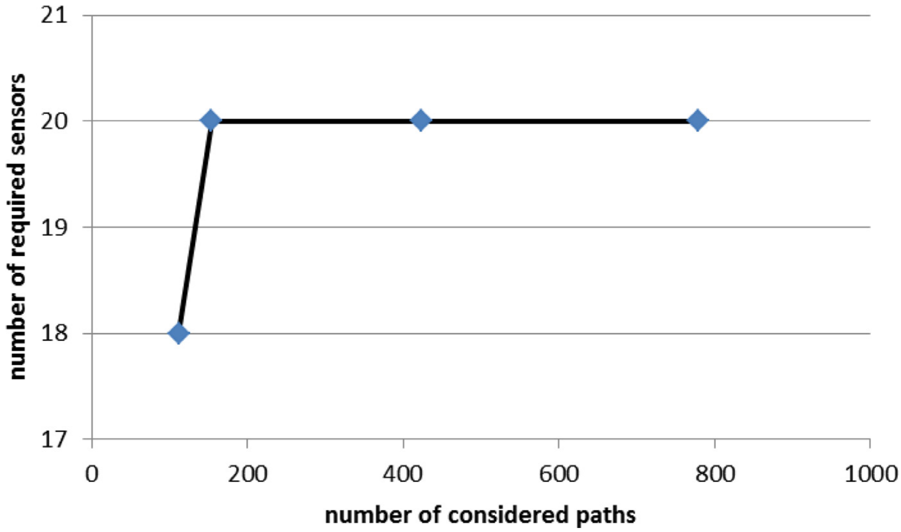

Fig. 14. Optimal solutions of model L_3 for cases 1–4 of example 4.

In this constraint, the value inside the parentheses has four indices: a, a', r and r'. In real examples, this value is equal to one for only a very small percentage of all possible quadruples and is zero for the rest. For example, in the third case of the current example, the total number of quadruples is $76^2 \times 424^2 = 1.04 \times 10^9$. There are only 216 quadruples, or 2×10^{-5}%, for which the value in parentheses is non-zero. Thus, in real problems and for large networks, it is preferable to build this constraint based on only these few sets of indices and not the entire matrix.

4.3.2. Model L_3

Model L_3 yields the minimum number of sensors required for full observability of the OD flows; the results of solving this model for the four cases of example 4 are reported in this section. Fig. 14 shows the number of sensors required to guarantee the uniqueness of the OD vector. Fig. 15 shows the CPU time required to solve the model for each case as a function of the number of routes.

Fig. 14 illustrates an interesting aspect of model L_3. Lemma 4 states that a solution to model L_3 always exists. An obvious solution to this model is that in which every arc of the network has a sensor. However, there is another feasible solution to this model that contains far fewer sensors, and its target function is much less.

Lemma 6. *Consider the sensor location vector in which all outgoing arcs from all origins and all incoming arcs to all destinations are equipped with a vehicle identification sensor. For this location vector, the set of OD vectors that satisfy condition (2) is single-membered.*

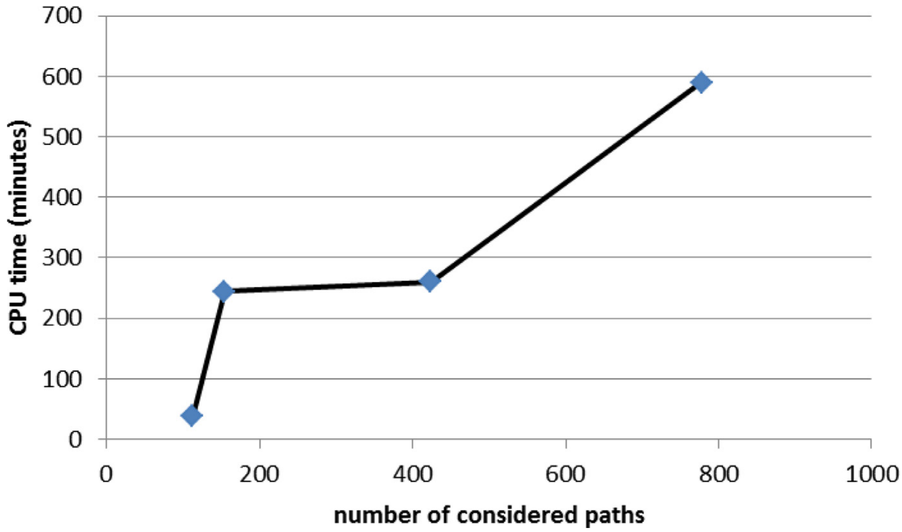

Fig. 15. CPU times required to solve model $\boldsymbol{L}_3$ for cases 1–4 of example 4.

The proof of Lemma 6 is trivial and is omitted here. Lemma 6 presents a new solution that is feasible for model L_3 independent of the set of routes considered. Because model L_3 minimizes its target function, this new solution yields an upper bound. Moreover, as the number of routes considered for each OD pair is increased, the solution of model L_3 converges to this solution. Obviously, when the solution reaches this point, further increasing the number of considered routes for each OD pair, will no longer increase the number of required sensors. This new feasible solution can be very helpful in solving model L_3 for large networks.

For the four cases of example 4, this new solution for model L_3 contains 40 sensors. However, the number of required sensors remains constant at 20 after the second case, in which 154 routes are considered. Unlike Lemma 6, this behavior is not a global aspect of model L_3; instead, it is related to the set of considered OD pairs. In what follows, the reason for this saturation will be discussed. In this network, nodes 1, 2, 3, 6, 7, 8, 12, 13, 18, 20, 21 and 24 constitute 12 OD pairs. Each of these nodes is the origin in exactly one OD pair and is the destination in exactly one other OD pair. Consider the following set of arcs:

$$\{1, 2, 3, 4, 5, 6, 7, 8, 14, 33, 35, 36, 37, 38, 39, 66, 73, 74, 75, 76\}$$

This set contains all outgoing and incoming arcs of nodes 1, 2, 3, 12, 13 and 24. These nodes are the origin nodes in 6 of the considered OD pairs and are the destination nodes in the other six OD pairs. When all incoming and outgoing arcs of a node A are equipped with a vehicle identification sensor, one can distinguish between the following four sets of routes:

1. Routes that start from node A
2. Routes that end at node A
3. Routes that pass through node A, but this node is not their origin or destination
4. Routes that do not contain node A

For example, consider node 1. There are two ingoing arcs, {3, 5}, and two outgoing arcs, {1, 2}. The routes in the first set correspond to combinations that contain either arc 1 or arc 2 and do not contain arc 3 and arc 5. The routes in the second set correspond to combinations that contain either arc 3 or arc 5 and do not contain arc 1 and arc 2. The routes in the third set correspond to combinations that contain one of the following combinations: (3, 1), (3, 2), (5, 1) and (5, 2). Finally, the routes in the fourth set correspond to combinations that do not contain any of arcs 1, 2, 3 and 5. Moreover, in this example and for this solution, every route has at least one sensor. Thus, each route corresponds to a non-empty combination. Therefore, based on each node, the set of all routes can be partitioned into four subsets. Based on this partitioning, the routes belonging to each OD pair can be distinguished from the routes belonging to other OD pairs. This is the necessary and sufficient condition for the uniqueness of the OD pairs. Thus, in example 4, as the number of routes between each OD pair arbitrarily grows, the size of the solution to model L_3 will remain fixed at 20 sensors.

4.4. Example 5, a real case study: Mashhad network

In this example, the topology of the Mashhad network is considered. This network was divided into 141 traffic zones and contains 2526 arcs, 917 nodes and 7157 OD pairs with non-zero demands (Poorzahedi et al., 1997). In this study “an origin–destination survey was conducted from 4% of the population by house interviewing method and was validated by observations from several screen lines in the study area. Validation of the OD flows is achieved by the actual traffic counts

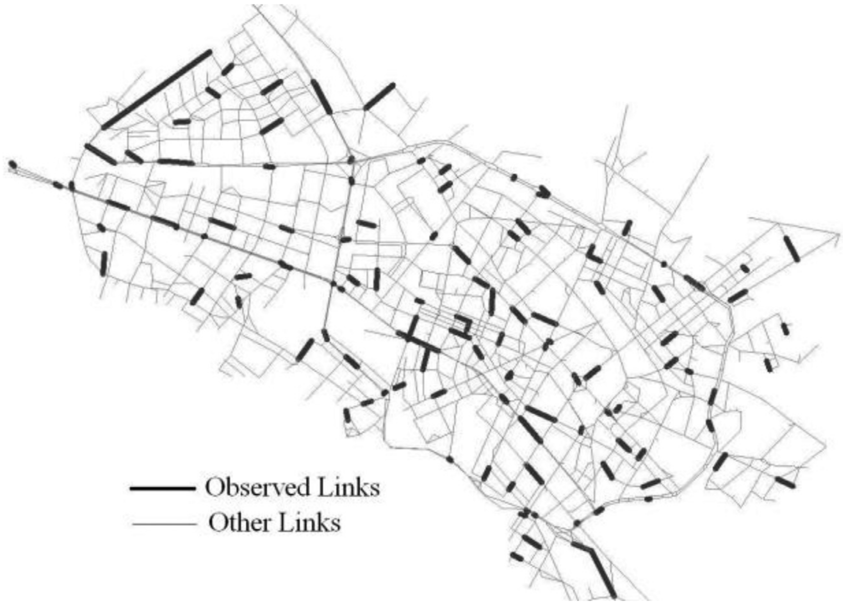

Fig. 16. Mashhad city network (Poorzahedi et al., 1997).

done simultaneously on 112 links of the network during peak hour. Fig. 16 shows the Mashhad city network and the observed links" (Poorzahedi et al., 1997). At the time of the cited study, Mashhad had a population of approximately 2 million. In that study, Poorzahedi and coauthors assigned the surveyed OD vector to the network based on the user equilibrium assumption using the complementary algorithm (Aashtiani, 1979). In their results, there were 7503 paths with non-zero flows; thus, there were 1.05 paths on average per OD pair.

In this example, these OD pairs have been sorted based on their demand, from highest to lowest. Model L_1 was solved for three sets of routes: the routes belonging to the first 40 OD pairs, the routes belonging to the first 50 OD pairs and the routes belonging to the first 60 OD pairs.[6] The percentages of the demand covered in cases 1–3 are 5.3, 6.2 and 7.1%, respectively. The total numbers of routes are 48, 59 and 69 for cases–3, respectively. The total numbers of arcs used in at least one of the considered routes are 362, 402 and 430 for cases 1–3, respectively. The main aspect of this example that makes it a medium sized real case study is its high number of used arcs. The three obtained solutions contain 37, 46 and 52 sensors for cases 1–3, respectively. The required CPU times for these three examples are 37.7, 180.5 and 282.6 min respectively.

The results of our numerical study, particularly those for this example and example 4, illustrate the amounts of time required to solve our location models using a general solver such as CPLEX. It is apparent from these examples that solving these models for even medium-sized networks is a time-consuming task. This demonstrates the need to develop model-specific algorithms for solving these models for large networks. This will be an interesting topic for future research. As an alternative to developing model-specific algorithms, another way to solve these models for larger examples is to control the size of the problem. There are two important factors that define the complexity of solving models L_1, L_2, L_3 and L_4 for a given network. The first of these factors is the set of considered paths, and the second is the set of possible arcs on which sensors can be located. The set of OD pairs also is a major contributor to the complexity of models L_3 and L_4. For a large network, the size of the problem can be reduced by modifying these three values. The mathematical relations between these values and the complexity of these four models will be another interesting topic for future research. In the following, the particular case of modifying the set of routes to reduce the size of the problem will be discussed.

4.4.1. Set of routes

In each location model presented in this paper, as in model C, the set of paths is a key parameter in defining the complexity of the model and its solution. For example, model L_1 contains $2r^2 + 4r + \frac{3}{2}(n^2 + n)$ constraints in addition to the integrality condition, where r is the number of paths and n is the number of arcs. There is also a more compact version of this model that contains $\frac{1}{2}r^2 + \frac{3}{2}r + \frac{3}{2}(n^2 + n)$ constraints. Model C contains $\frac{1}{2}(r^2 + r)$ constraints. Considering only the important paths has always been a common approach adopted by researchers; for example, one might consider only the paths that had positive flows in past studies. In the current example, example 5, this approach was applied to the Mashhad network. In this example, the set of paths to be considered was defined based on the paths that had non-zero flows

[6] Here, the first 40 OD pairs are the 40 OD pairs with the highest demand.

after a user-equilibrium-based assignment of the OD vector. Another approach would be to also consider paths with similar qualities to those with non-zero flows.

The dependence of the time required to solve model L_1 on the number of considered paths is shown in the results of examples 4 and 5. In both cases, the CPU time grows very rapidly as the number of paths increases. However, the numbers of required sensors in models L_1 and L_3 exhibit a milder rate of growth with an increasing number of considered paths.

When the target flow is the OD vector, the results of example 4 for model L_3 and Lemma 6 together reveal an interesting fact: as the number of considered paths increases, the size of the solution to model L_3 will eventually reach saturation. In other words, the optimal solution will not always continue to grow with an increase in the number of considered paths. Another observation that is promising for reducing the number of considered paths is that in the Mashhad network example, there were found to be 1.05 paths on average per OD pair. If one accepts equilibrium as a reasonable assumption, especially considering the growing market penetration of GPS devices and smart phones, this fact can be used to reduce the number of considered paths.

5. Conclusions

In this paper, several of the existing location models were discussed. To our knowledge, none of the existing location models for vehicle identification sensors considers the order in which a vehicle is detected by the sensors, although this order is available in the data. When the sensor order is taken into account, the existing models no longer yield the minimum number of sensors required for full observability of the target flows. Therefore, in this paper, four location models that consider the order of the sensors were presented and proven correct. The definitions of these four models are as follows:

- Model L_1: This model minimizes the number of sensors required to ensure the uniqueness of the path flows when the sensor order is available in the data.
- Model L_2: This model maximizes the number of fully observable path flows subject to a budget constraint when the sensor order is available in the data.
- Model L_3: This model minimizes the number of sensors required to ensure the uniqueness of the OD flows when the sensor order is available in the data.
- Model L_4: This model maximizes the number of fully observable OD flows subject to a budget constraint when the sensor order is available in the data.

For several numerical examples, these four models were solved using the GAMS software. These numerical examples include several medium-sized examples, including an example of a real-world large-scale transportation network in Mashhad. As the results show, solving model L_1 for some of our examples required several days of CPU time. Although the examples solved in this paper include several medium-scale cases of large networks, solving these models for larger networks remains an interesting area for future research.

References

Aashtiani, H.Z., 1979. The Multi-modal Traffic Assignment Problem. Massachusetts Institute of Technology.

Bianco, L., Confessore, G., Gentili, M., 2006. Combinatorial aspects of the sensor location problem. Annals of Operations Research 144 (1), 201–234.

Bierlaire, M., 2002. The total demand scale: a new measure of quality for static and dynamic origin–destination trip tables. Transportation Research Part B: Methodological 36 (9), 837–850.

Castillo, E., Calvino, A., Menéndez, J.M., Jimenez, P., Rivas, A., 2013. Deriving the upper bound of the number of sensors required to know all link flows in a traffic network. IEEE Transactions on Intelligent Transportation Systems 14 (2), 761–771.

Castillo, E., Gallego, I., Menéndez, J.M., Rivas, A., 2010. Optimal use of plate-scanning resources for route flow estimation in traffic networks. IEEE Transactions on Intelligent Transportation Systems 11 (2), 380–391.

Castillo, E., Jiménez, P., Menéndez, J.M., Conejo, A.J., 2008. The observability problem in traffic models: algebraic and topological methods. IEEE Transactions on Intelligent Transportation Systems 9 (2), 275–287.

Castillo, E., Menéndez, J.M., Jiménez, P., 2008. Trip matrix and path flow reconstruction and estimation based on plate scanning and link observations. Transportation Research Part B: Methodological 42 (5), 455–481.

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.

Doblas, J., Benitez, F.G., 2005. An approach to estimating and updating origin–destination matrices based upon traffic counts preserving the prior structure of a survey matrix. Transportation Research Part B: Methodological 39 (7), 565–591.

Ehlert, A., Bell, M.G., Grosso, S., 2006. The optimisation of traffic count locations in road networks. Transportation Research Part B: Methodological 40 (6), 460–479.

Eisenman, S., Fei, X., Zhou, X., Mahmassani, H., 2006. Number and location of sensors for real-time network traffic estimation and prediction: sensitivity analysis. Transportation Research Record: Journal of the Transportation Research Board 1964, 253–259.

Gan, L., Yang, H., Wong, S.C., 2005. Traffic counting location and error bound in origin–destination matrix estimation problems. Journal of Transportation Engineering 131 (7), 524–534.

Gentili, M., Mirchandani, P., 2012. Locating sensors on traffic networks: models, challenges and research opportunities. Transportation Research Part C: Emerging Technologies 24, 227–255.

Gentili, M., Mirchandani, P.B., 2005. Locating active sensors on traffic networks. Annals of Operations Research 136 (1), 229–257.

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), 873–894.

Li, X., Ouyang, Y., 2011. Reliable sensor deployment for network traffic surveillance. Transportation Research Part B: Methodological 45 (1), 218–231.

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.

Mirchandani, P.B., Gentili, M., He, Y., 2009. Location of vehicle identification sensors to monitor travel-time performance. IET Intelligent Transport Systems 3 (3), 289–303.

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.

Ng, M., 2013. Partial link flow observability in the presence of initial sensors: solution without path enumeration. Transportation Research Part E: Logistics and Transportation Review 51, 62–66.

Poorzahedi, H., Kermanshah, M., Aashtiani, H.Z., Shafahi, Y., 1997. Traffic Assignment Model and Mashhad Transportation Network Performance in 1994. Sharif University of Technology, Institute of Transportation Studies and Research submitted to Mashhad Municipality.

Sherali, H.D., Desai, J., Rakha, H., 2006. A discrete optimization approach for locating automatic vehicle identification readers for the provision of roadway travel times. Transportation Research Part B: Methodological 40 (10), 857–871.

Viti, F., Rinaldi, M., Corman, F., Tampère, C.M., 2014. Assessing partial observability in network sensor location problems. Transportation Research Part B: Methodological 70, 65–89.

Viti, F., Verbeke, W., Tampère, C., 2008. Sensor locations for reliable travel time prediction and dynamic management of traffic networks. Transportation Research Record: Journal of the Transportation Research Board 2049, 103–110.

Yang, H., Iida, Y., Sasaki, T., 1991. An analysis of the reliability of an origin–destination trip matrix estimated from traffic counts. Transportation Research Part B: Methodological 25 (5), 351–363.

Yang, H., Zhou, J., 1998. Optimal traffic counting locations for origin–destination matrix estimation. Transportation Research Part B: Methodological 32 (2), 109–126.

Zhou, X., List, G.F., 2010. An information-theoretic sensor location model for traffic origin–destination demand estimation applications. Transportation Science 44 (2), 254–273.