Research Article

Sensor Location Problem for Network Traffic Flow Derivation Based on Turning Ratios at Intersection

Minhua Shao, Lijun Sun, and Xianzhi Shao

1 Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 200092, China
2 China Airport Construction Group Corporation, Beijing 100101, China

Correspondence should be addressed to Minhua Shao; shaominhua@tongji.edu.cn

Received 28 August 2015; Accepted 28 January 2016

Academic Editor: Luca D’Acierno

Copyright © 2016 Minhua Shao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The sensor location problem (SLP) discussed in this paper is to find the minimum number and optimum locations of the flow counting points in the road network so that the traffic flows over the whole network can be inferred uniquely. Flow conservation system at intersection is formulated firstly using the turning ratios as the prior information. Then the coefficient matrix of the flow conservation system is proved to be nonsingular. Based on that, the minimal number of counting points is determined to be the total number of exclusive incoming roads and dummy roads, which are added to the network to represent the trips generated on real roads. So the task of SLP model based on turning ratios is just to determine the optimal sensor locations. The following analysis in this paper shows that placing sensors on all the exclusive incoming roads and dummy roads can always generate a unique network flow vector for any network topology. After that, a detection set composed of only real roads is proven to exist from the view of feasibility in reality. Finally, considering the roads importance and cost of the sensors, a weighted SLP model is formulated to find the optimal detection set. The greedy algorithm is proven to be able to provide the optimal solution for the proposed weighted SLP model.

1. Introduction

Link flow data in the road network is the valuable information in many applications of traffic planning and management, such as the road network planning and the congestion analysis. Placing traffic sensors on every link is usually not realistic, because there are thousands of links in the urban road network normally and the investment on sensors installation and maintenance will be huge. So it is important to identify the subset of links so that the traffic flow on every link in the network can be inferred if sensors are installed on this subset of links to collect flow data. The sensor location problem (SLP) was proposed to address this problem.

The SLP has attracted increasing interest in the past few years for its importance in the area of transportation system analysis. Two main classes of SLP were summarized by Gentili and Mirchandani [1]: sensor location flow-observability problems and sensor location flow-estimation problems. The objective of the first-class SLP is to find the optimal sensor locations so that all link flows (or partial link flows) can be determined uniquely. The second-class SLP is trying to improve the accuracy of the estimation by optimizing the sensors locations. Early research in the literature mainly focused on the second-class SLP as a subproblem of origin-destination (OD) estimation. The “OD Covering Rule” was given by Yang et al. [2, 3] to guarantee the upper-bounded OD estimation error. Several other rules were also proposed to improve the estimation accuracy, like the “maximal flow fraction rule,” the “maximal flow-intercepting rule,” and so on (for details, see [1–5]). The first-class SLP was not formally defined until 2001. Bianco et al. [6] defined SLP as determining the minimum number and locations of counting points to infer all traffic flows in the network. We will focus on the first-class SLP in this paper.

The first-class SLP was proposed in the context of traffic planning and also was developed in this domain in most of the following research. So, essentially, it is a static optimization problem addressing the method of network sensor deployment under the steady-state traffic condition. Generally, in the sensor location flow-observability problems,
a system of linear equations will be formulated at first, which represents the relationship of different link flows or the relationship between link flows and route flows (or OD trip) under the law of flow conservation. This system of equations is defined as the flow conservation system in this paper. Then, by deploying the sensors, equations representing the detected link flows will be added. These equations are defined as the detecting equations. The number of the detecting equations is equal to the number of sensors. Basically, the main idea of the first-class SLP is to find the optimal detecting equations so that all the variables are solvable. According to this idea, the SLP was normally designed as an integer linear programming model in the existing studies. The difference among these studies lies in the assumptions on the prior information, which lead to different formulations of the flow conservation system. According to the prior information, three kinds of flow conservation systems are formulated: OD-link based system, route-link based system, and link-link based system [1]. The OD-link based system uses the link choice proportions as the prior information, which is defined as the proportion of trips between one OD pair that uses a given link in the network. Similarly, the route-link based system uses the link–route coefficients as the prior information. When a link is used by a route, the corresponding coefficient will be 1; otherwise, the coefficient will be zero [7]. These two kinds of systems use the OD trips or route flows as the bridge between detected link flows and undetected link flows. So the OD matrix and route flows will be obtained at the same time. However, in the scenarios when the link flow is the only information we care about, the effort used in calculating the OD trips and route flows is not necessary. In particular for the route-based system, enumerating all routes is not easy for a large-scale network. At the same time, if we are only interested in the link flows of one small local network, the whole network still needs to be considered in the SLP since the OD trip and route flow are defined over the network. The third kind of system uses the relationship of the link flows directly. Without using the concept of OD trip or route flow, this kind of system can be used to formulate the SLP for local network. Bianco et al. [6] proposed the first system of this kind in 2001, using the split ratios of outgoing roads at intersections as the prior information. The split ratio of an outgoing link at an intersection is defined as the fraction of the flow on the outgoing link over the total incoming flow of the intersection. With the assumption of symmetrical network and node-based detection (the sensors are located on every in- and out-link of the intersection), a necessary condition was given to guarantee that the number of equations in the system is not less than the number of variables. They also presented a couple of greedy heuristics to find the lower and upper bounds on the number of sensors for random networks. After that, Bianco et al. [8] proved that the SLP proposed in 2001 is NP-hard in the computation complexity. Linear algorithms were invented for several special graphs, including the paths, circles, and combs. Based on the work of Bianco, Morrison et al. [9, 10] gave a stricter necessary condition for general graphs and proved the condition to be also sufficient for the special network with trees as its unmonitored components. Essentially, the network structure-driven logic is used in the link-link based sensor location flow-observability problems. So the traffic flows over the network are calculated based on the network topology and any behavior assumption of the road users is not required [6]. Until now, most discussions on the link-link based sensor location problems used the split ratios as the prior information. The network was assumed to be symmetrical and the detection was assumed to be node based. However, in reality, unidirectional roads normally exist, which leads to an asymmetrical network. At the same time, road-based detection is more reasonable. In this paper, the network is allowed to be asymmetrical and the detection is assumed to be road based. Turning ratios at intersections are used as the prior information. The turning ratio is defined as the ratios of flow turning from an incoming link to an outgoing link over the total flow of the incoming link. Turning ratio is used as a popular parameter in many applications. Many methods were designed to obtain the turning ratios (for details, see [11–13]), which make the estimation of turning ratios from historical data possible. Using turning ratios as the prior information, an integer linear programming model is designed in this paper to address the sensor location flow-observability problem under steady-state traffic conditions.

The present paper is organized as follows. In Section 2, the definition and assumptions of the road network used in this paper are specified. Then, the flow conservation system is formulated in Section 3 based on the turning ratios. After that, the coefficient matrix of the flow conservation system is proven to be full rank in row space. Based on that, in Section 5, the SLP model is given and the greedy algorithm is set up to solve the model. Finally, some conclusions are summarized.

2. Definitions and Assumptions

Suppose the road network is composed of $n$ directed roads. The roads are numbered by $1, 2, \ldots, n$. Represent the roads set as $\text{RD} = \{1, 2, \ldots, n\}$. Six assumptions are used in this paper.

Assumption 1. The road network which we are interested in is taken from a bigger one so that on the boundary there are directed roads whose upstream or downstream intersections are missing.

Definitions for Three Categories of Roads. Under Assumption 1, there must be three categories of roads in a network: incoming/outgoing road, which connects two adjacent intersections; exclusive incoming road, whose downstream intersection is in the network but the upstream intersection is not; exclusive outgoing road, whose upstream intersection is in the network but the downstream intersection is not. Figure 1 gives an example of a simple network including only two intersections.

In the network of Figure 1, roads 4 and 8 are the incoming/outgoing roads, roads 9, 10, 11, 12, 13, and 14 are the exclusive incoming roads, and roads 1, 2, 3, 5, 6, and 7 are the exclusive outgoing roads.
Mathematical Problems in Engineering

The roads compose set \( \mathcal{F} = \{n+1, n+2, \ldots, n+g\} \).

Assumption 2. Turning ratios for any road in \( \mathcal{F} \) are known.

**Definition for Turning Ratios.** The turning ratio is defined as the proportion of the flow on an incoming road which turns to an outgoing road at the intersection. Denote the turning ratio of traffic flow turning from road \( j \) to road \( i \) at intersection as \( a_{ij} \) and the traffic flow on road \( j \) as \( x_j \). The traffic flow turning from road \( j \) to road \( i \) at intersection is \( a_{ij} \times x_j \).

Assumption 3. Flow conservation law is satisfied at the intersection.

For an outgoing road \( i \) without trips originating or terminating, the traffic flow is equal to the sum of the flows turning into it from all the incoming roads at the intersection. Using \( x_i \) as the traffic flow on outgoing road \( i \), we have

\[
x_i = \sum_{j \in \mathcal{F}_{\text{out}}, j \neq i} a_{ij} x_j.
\]

**Definition for Net Balancing Flow.** For the outgoing roads which have trips originating or terminating, the net balancing flow of the outgoing roads is defined as the originating trips minus the terminating trips. Denote the net balancing flow on road \( k \) as \( s_k \) and then for road \( k \), we have

\[
x_k = \sum_{j \in \mathcal{F}_{\text{out}}, j \neq k} a_{kj} x_j + s_k.
\]

Suppose there are \( g \) roads in set \( \mathcal{F}_{\text{out}} \) which have nonzero net balancing flows. Normally, \( g \) is far smaller than the size of \( \mathcal{F}_{\text{out}} \), especially when the analysis period is long enough.

Assumption 4. For any road set \( A \subseteq \mathcal{F}_{\text{in}} \), there must be at least one road \( j \in A \) which has a turning ratio \( a_{ij} > 0 \), where road \( i \notin A \).

If in road set \( A \subseteq \mathcal{F}_{\text{in}} \), all the turning ratios are directed to the roads inside the set, and then the traffic inside the set will never flow out. So Assumption 4 is actually to say that the turning ratios should not define a self-circulation system. This assumption is reasonable because any part of the road network has to be able to exchange traffic with another part.

Assumption 5. Traffic flows are assumed to be detected based on roads.

Once a road is detected, the traffic flow on the road is known.

Assumption 6. The detecting flows and turning ratios are both assumed to be error-free.

3. Flow Conservation System Based on the Turning Ratios

From Assumption 3 in Section 2, we know that, for any road in set \( \mathcal{F}_{\text{out}} \), either (1) or (2) can be formulated. In order to simplify the expressions of the flow conservation system and the corresponding parameters, we renew \( \mathcal{F} \) by numbering all the roads in \( \mathcal{F}_{\text{out}} \) at first so that \( \mathcal{F}_{\text{out}} = \{1, 2, \ldots, m\} \) and \( \mathcal{F}_{\text{in}} = \{m+1, m+2, \ldots, n\} \). On the other hand, the nonzero net balancing variable \( s_i \) is looked at as the traffic flow on a dummy road which is an incoming road of the corresponding outgoing road with 1 as the turning ratio. The dummy roads can be looked at as the exclusive incoming roads. Compared to the real exclusive incoming roads, the dummy roads may have negative flows. To illustrate the method to add the dummy roads into the network, we suppose that there is a nonzero net balancing flow on road 4 in the network of Figure 1. Dummy road 15 is added to the network as shown in Figure 2. The turning ratio from dummy road 15 to road 4 is set to be 1, which means all the flow on dummy road 15 will turn into road 4.

The dummy roads are numbered after exclusive incoming roads. The \( g \) roads compose set \( F = \{n+1, n+2, \ldots, n+g\} \).
Using $X = [x_1, x_2, \ldots, x_j, \ldots, x_{m+g}]'$ to represent the traffic flow vector of the network, where $x_j$ is the traffic flow on road $j$, $j \in \text{RD}_\text{in} \cup \text{F}$, the flow conservation system can be formulated as follows:

$$
\sum_{j \in \text{RD}_\text{in} \cup \text{F}, j \neq 1} a_{1j} x_j - x_1 = 0,
$$

$$
\sum_{j \in \text{RD}_\text{in} \cup \text{F}, j \neq 2} a_{2j} x_j - x_2 = 0,
$$

$$
\vdots
$$

$$
\sum_{j \in \text{RD}_\text{in} \cup \text{F}, j \neq m} a_{mj} x_j - x_m = 0.
$$

(3)

The coefficient matrix $C_1$ of system (3) can be expressed as follows:

$$
C_1 = \begin{pmatrix}
  a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1m} & \cdots & a_{1(n+g)} \\
  a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2m} & \cdots & a_{2(n+g)} \\
  \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \cdots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} & \cdots & a_{m(n+g)}
\end{pmatrix}.
$$

(4)

$C_1$ is $m \times (n+g)$ matrix. The entries $a_{ij}$ are given as follows:

$$
a_{ij} = \begin{cases} 
-1 & \text{when } i = j \\
0 & \text{when } i \neq j \text{ and road } j \text{ has flows turning onto road } i \text{ at an intersection} \\
0 & \text{otherwise}.
\end{cases}
$$

(5)

The column vector of matrix $C_1$ is denoted as $a_j$, $a_j = [a_{1j}, a_{2j}, \ldots, a_{mj}]'$. $C_1$ can be written as $C_1 = (a_1, a_2, \ldots, a_{m+g})$. $C_2$ can be written as $C_2 = (a_1, a_2, \ldots, a_m)$. The flow conservation system can be rewritten as follows:

$$
\sum_{j=1}^{m} a_j x_j = 0 - \sum_{k=m+1}^{n+g} a_k x_k.
$$

(8)

4. Proof for the Full Row Rank of Coefficient Matrix

The objective of sensor location flow-observability problem is to find the minimum number and locations of counting points to infer all traffic flows in the network. If we can prove that the coefficient matrix $C_1$ of the flow conservation system is full rank in the row space, then the minimum number of counting points will be determined. The task of corresponding SLP will be reduced to optimize the sensors location only.

According to the knowledge of linear algebra, if we can find a nonsingular $m \times m$ sub matrix in $C_1$, then matrix $C_1$ is proven to be full rank in the row space.

Let us select the first $m$ columns from $C_1$ to compose a square matrix $C_2$:

$$
C_2 = \begin{pmatrix}
  a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1m} \\
  a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2m} \\
  \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn}
\end{pmatrix}.
$$

(7)

The variables $x_j$, $j \in \text{RD}_\text{out}$, on the left side correspond to the traffic flows on outgoing roads. On the right side, $x_k$, $k \in \text{RD}_\text{in}$, correspond to the traffic flows on incoming roads; $x_k$, $k \in F$, correspond to the net balancing flows.

If we can prove that the traffic flows on the outgoing roads must be zero when the right side vector of (8) equals $0$, then the column vectors are proven to be linear independent, which means $C_2$ must be nonsingular.

Proposition 7. If the right side vector of (8) equals $0$, then the traffic flows on all the exclusive outgoing roads must be zero.

Proof. According to the analysis above, we have

$$
\sum_{1 \leq j \leq m} a_{ij} = 1, \quad j \in \text{RD}_\text{in} \cup \text{RD}_\text{out} \cup \text{F}
$$

$$
\sum_{1 \leq j \leq m} a_{ij} = 0, \quad j \in \text{RD}_\text{out}
$$

(9)

$$
a_{ij} = -1, \quad i = j.
$$

Adding all the equations in system (8) together, we will easily get that the sum of the left side of (8) is minus total traffic flows in set $\text{RD}_\text{out}$ and the sum of the right side of (8) is minus total traffic flows in set $\text{RD}_\text{in} \cup \text{F}$. Set the right side vector of (8) $0$. So total traffic flows in set $\text{RD}_\text{in} \cup \text{F}$ are zero. Then, total traffic flows in set $\text{RD}_\text{out}$ are zero too,
which determines that all the exclusive outgoing roads have no traffic on them.

Actually, this is an obvious result under the assumption of traffic flow conservation. For a road network, all the traffic flows (including the net balancing flows here) entering the network must equal the traffic flows leaving the network.

**Proposition 8.** If the right side vector of (8) equals 0, then the traffic flows on all the incoming/outgoing roads must be zero too.

**Proof.** Proof by contradiction is used here to prove Proposition 8.

Suppose there are incoming/outgoing roads which have nonzero traffic flows on them. All these roads compose a road set $A'$. We have $A' \subset R_{D_{in}}$. According to Assumption 4 in Section 2, there must be at least one road $j \in A'$ which has a turning ratio $a_{ij} > 0$, where road $i \notin A'$. Since the traffic flow on road $j$ is not zero, there must be $a_{ij} \times x_i$, traffic flow turning from road $j$ into road $i$. Then, the traffic flows on road $i$ must not be zero. This can be proved as follows.

For any road $i$ in set $R_{D_{out}}$, the incoming traffic flows at the intersection can be divided into two parts: traffic flows coming from the exclusive incoming roads or/dummy roads which correspond to the right side of (8) and traffic flows coming from the incoming/outgoing roads. So for a road $i$ in set $R_{D_{out}}$, if the right side of (8) is set to zero and there are traffic flows turning from the incoming/outgoing roads into it, the traffic flows on it must be nonzero.

If road $i$ is still an incoming/outgoing road, set $A' = A' \cup \{i\}$. Then, we can still find at least one outgoing road outside $A'$, whose traffic flow is not zero. Since $R_{D_{out}} = R_{D_{in}} \cup R_{D_{out}}$, there must be at least one exclusive outgoing road which has nonzero flow. This cannot happen according to Proposition 7. So if the right side vector of (8) equals 0, then the traffic flows on all the incoming/outgoing roads must be zero too.

Now we have proven that if the right side vector of (8) equals 0, the traffic flows on the outgoing roads must be zero. So $C_2$ is nonsingular, which means that the coefficient matrix $C_1$ of the flow conservation system is full rank in the row space.

### 5. Sensor Location Optimization Model and Solution Algorithm

#### 5.1. Full Flow-Observability SLP Based on Turning Ratios

**5.1.1. SLP Description.** As we stated above, the objective of full flow-observability SLP is to determine the minimum number and locations of the sensors over the network so that all the traffic flows can be inferred.

Based on turning ratios, $m$ linear equations can be formulated to compose the flow conservation system as system (6). The number of variables in the system is $n + g$, which is larger than $m$. It is easy to know that at least other $n + g - m$ equations are needed to derive all the variables uniquely. By installing sensors on road, detecting equation can be formulated as $x_j = b_j$, where $b_j$ is the detected traffic flow. So at least $n + g - m$ roads need to be detected. Denote the coefficient matrix for the $n + g - m$ detecting equations as $C_3$. Matrix $C_3$ is a $(n + g - m) \times (n + g)$ matrix where only one entry equals one and the others equal zero in each row. For the undetected roads, the corresponding column vectors are zero vectors. Adding all the detecting equations, the complete system can be formulated as follows:

$$
(\begin{bmatrix} C_1 \\ C_3 \end{bmatrix}) \times \begin{pmatrix} X \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ B \end{pmatrix},
$$

where $B$ is the detected flows vector, composed of the detected traffic flows $b_j$.

System (10) will generate a unique solution if and only if the coefficient matrix is nonsingular. So the objective of full flow-observability SLP here is to find a detecting coefficient matrix $C_3$ so that the coefficient matrix of system (10) is nonsingular.

In Section 4, matrix $C_3$ is proven to have full row rank. So we can always find a matrix $C_3$ with $n + g - m$ to generate a nonsingular coefficient matrix for system (10), which means only $n + g - m$ roads are required to be detected.

**5.1.2. SLP Modeling.** Using $\delta_j$ as the decision variable in the model to determine whether road $j$ will be detected or not, $\delta_j = 1$ means road $j$ will be detected; $\delta_j = 0$ means road $j$ will not be detected. We design a matrix $C$ as follows:

$$
C = \begin{pmatrix}
    a_{i1} & a_{i2} & \cdots & a_{im} & \cdots & a_{i(n+g)} \\
    a_{21} & a_{22} & \cdots & a_{2m} & \cdots & a_{2(n+g)} \\
    \vdots & \ddots & & \vdots & & \vdots \\
    a_{n1} & a_{n2} & \cdots & a_{nm} & \cdots & a_{n(n+g)} \\
    \delta_1 & \delta_2 & \cdots & \delta_m \\
\end{pmatrix}
$$

Let $\delta = [\delta_1, \delta_2, \ldots, \delta_j, \ldots, \delta_m]^T$. The full flow-observability SLP based on turning ratios can be described as follows.

**Determine $\delta$ which satisfy**

$$
\text{rank}(C) = n + g
$$

$$
\sum_{j=1}^{n+g} \delta_j = n + g - m,
$$

where $\text{rank}(C)$ is the rank of matrix $C$.

The solutions of model (12) must exist according to the analysis in Section 5.1.1.
5.1.3. A Special Design: Closing Ring. Shao et al. [14, 15] introduced a special design of sensor location over the network called “closing ring,” which is given in Figure 3.

The “closing ring” is designed to install counting sensors on every exclusive incoming road of the network with the assumption that no net balancing flows exist. No proof was given in [14, 15] to guarantee a unique flow solution. By the analysis above, the assumption can be released by adding dummy roads into the network to represent the net balancing flows. So here the “closing ring” is designed to detect the flows on all exclusive incoming roads and dummy roads. This design can be easily proven to be a solution of model (12).

In $C_1$, the last $n + g - m$ columns correspond to the exclusive incoming roads and dummy roads of the network. So the corresponding coefficient matrix for the detecting equations of closing ring design is $C_3 = [O \ I]$, where $I$ is an $(n + g - m) \times (n + g - m)$ identity square matrix and $O$ is $(n + g - m) \times m$ zero matrix. The whole coefficient matrix $C_f$ of system (10) can be written as follows:

$$C_f = \begin{pmatrix} C_2 & C'' \end{pmatrix} \begin{pmatrix} O & I \end{pmatrix},$$

(13)

where $C''$ is the submatrix composed of the last $n + g - m$ columns of $C_1$.

Obviously, the determinant of $C_f$ can be calculated as 
$$\det(C_f) = \det(C_2) \times \det(I) = \det(C_2).$$

Since $C_2$ has been proven to be nonsingular, we have $\det(C_f) \neq 0$. So we have $\text{rank}(C_f) = n + g$. Obviously, $\sum_{j=1}^{n+g} j = n + g - m$ is satisfied.

We do not use any assumption of the network topology here. So, if possible, “closing ring” can always be a sensor location design for any road network, which can be given without any calculation.

5.1.4. Feasibility for Detecting Only Real Roads. In the “closing ring” design, all exclusive incoming roads and dummy roads are supposed to be detected. But, in reality, it is not easy to detect the net balancing flows. If we can always find $n + g - m$ real roads as a detection set in the above SLP, the solution will be more applicable in reality.

Renew $RD_{out}$ by numbering the roads with nonzero net balancing flows after others. It is equivalent to permute rows and first $m$ columns in matrix $C_1$. After that, permute the last $g$ columns in matrix $C_1$ so that the coefficient matrix for dummy roads is expressed by $C_4 = \left( \begin{smallmatrix} I_m \end{smallmatrix} \right)$, where $I_g$ is an $g \times g$ identity square matrix and $O$ is a $(m - g) \times g$ zero matrix. After all the permutations, the coefficient matrix $C_f^p$ of the flow conservation system is as follows:

$$C_f^p = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2m} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{(m-g)1} & a_{(m-g)2} & \cdots & a_{(m-g)m} & \cdots & a_{(m-g)m} \\ a_{(m-g+1)1} & a_{(m-g+1)2} & \cdots & a_{(m-g+1)m} & \cdots & a_{(m-g+1)m} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} & \cdots & a_{mm} \\ \end{pmatrix} I_g \quad (14)$$

Denote the first $m$ columns of $C_f^p$ as $C_2^p$. Submatrix $C_2^p$ is also the permutations result of matrix $C_2$. Since $C_2$ is nonsingular, $C_2^p$ is nonsingular too. So the row vectors in $C_2^p$ are linear independent. Consequently, the first $m - g$ row vectors are linear independent. We can always find a $(m-g) \times (m-g)$ nonsingular subsquare matrix $C_3$ in the first $m-g$ rows of $C_2^p$. Represent the submatrix, which is composed of $m-g$ columns in $C_2^p$ with $C_4$ as its first $m-g$ rows, as $C_6 = \left( \begin{smallmatrix} C_2 \ C_3 \end{smallmatrix} \right)$, where $C_5$ is the submatrix composed of the last $g$ rows of $C_6$. Also, represent the submatrix composed of the other $g$ columns of $C_2^p$ as $C_8$ and represent the submatrix composed of the columns from $(m+1)$th to $m$th in $C_f^p$ as $C_9$. By exchanging the columns in $C_8$ and $C_4$ and some columns permutations, $C_4^p$ can be written as $C_4^{pp}$:

$$C_4^{pp} = \left( C_6 \ C_4 \ C_9 \ C_8 \right). \quad (15)$$

As a subsquare matrix of $C_4^{pp}$, matrix $C_{10}$ = $(C_6 \ C_4)$ is a $m \times m$ matrix. We have

$$C_{10} = \begin{pmatrix} C_5 & O \\ C_7 & I_g \end{pmatrix}. \quad (16)$$

Since $C_6$ is a nonsingular square matrix, $I_g$ is an identity square matrix and matrix $C_1_{10}$ is also nonsingular. According to the analysis in “closing ring” design, if we place the sensors on all the roads corresponding to the columns in $C_8$ and $C_9$, the traffic flow vector can be inferred uniquely. Because all
the roads corresponding to the columns in $C_8$ and $C_9$ are real roads, we can always find $n + g - m$ real roads as a detection set for the SLP based on turning ratios.

5.2. Weighted Full Flow-Observability SLP Model Based on Turning Ratios

5.2.1. Model Formulation. Obviously, many combinations of roads in the network can be used as the solution detection sets for the full flow-observability SLP based on turning ratios. How to choose the detection set among all the feasible solutions depends on people’s evaluation on the importance of the roads and the cost of the sensors’ installation and maintenance. For example, an artery road is usually more important than a local street for the traffic management. So people may want to detect the traffic flow on it such that they can know its detailed traffic condition. Also lower cost of installation and maintenance could be preferred. A weighted vector can be given to the road by integrating people’s evaluation of installation and maintenance could be preferred. A weight vector could be given as $W = (\omega_1, \omega_2, \ldots, \omega_j, \ldots, \omega_{m+g})$, where $\omega_j$ is the weight of the road $j$. Here the dummy roads are included too. To get rid of dummy roads, we can simply set $\delta_j = 0$, $j \in \{n+1, \ldots, n+g\}$. Discussion in Section 5.1.4 shows that we can always find the feasible solutions even though $\delta_j = 0$, $j \in \{n+1, \ldots, n+g\}$.

The objective of the weighted SLP here is to find the optimal detection set so that the total weight of all the roads in the detection set is maximal. The weighted SLP model can be formulated as follows:

$$\begin{align*}
\text{max} & \quad W \times \delta \\
\text{s.t.} & \quad \text{rank}(C) = n + g \\
& \quad \sum_{j=1}^{m+g} \delta_j = n + g - m \quad (17) \\
& \quad \delta_j \in \{0, 1\}.
\end{align*}$$

All the variables and parameters have been defined before.

5.2.2. Solution Algorithm. Let us change $\delta_j$ in $C$ into 1; then we get $C'$:

$$
\begin{pmatrix}
\begin{array}{cccccccccccc}
a_{11} & a_{12} & \cdots & a_{1m} & \cdots & a_{1n} & \cdots & a_{1(n+g)} \\
a_{21} & a_{22} & \cdots & a_{2m} & \cdots & a_{2n} & \cdots & a_{2(n+g)} \\
\vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \cdots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn} & \cdots & a_{mn} & \cdots & a_{mn(n+g)} \\
1 & 1 & \cdots & 1 & \cdots & 1 & \cdots & 1
\end{array}
\end{pmatrix}
$$

Denote the $i$th row vector in $C'$ as $rw_i$. The first $m$ row vectors compose matrix $C_1$, $C'_1 = (rw_1', rw_2', \cdots, rw_m')$. The last $n + g$ row vectors compose identity matrix $I_{m+g}$, $I'_{m+g} = [rw_{m+1}' rw_{m+2}' \cdots rw_{m+n+g}']$. Every row vector $rw_{m+j}$ in $I_{m+g}$ is given a weight $\omega_j$. If $n + g - m$ row vectors of $I_{m+g}$ and $m$ row vectors in $C_1$ can compose a maximal independence row vector group of $C'$, then these $n + g - m$ row vectors are called feasible vector group.

Solving model (17) is essential to find the optimal feasible vector group whose total weight is maximal among all the feasible vector groups.

Greedy approach is used here to solve the SLP model (17). The algorithm is designed as follows.

Step 1. Set matrix $S = C_1$.

Step 2. Sort row vectors in $I_{m+g}$ into monotonically decreasing order by the given weight $\omega$.

Step 3. For each row vector $rw_{m+j}$, $j \in \{1, 2, \ldots, n+g\}$ in $I_{m+g}$, taken in monotonically decreasing order by weight $\omega_j$, consider the following.

Calculate the rank of $S \cup rw_{m+j}$ to see if $S \cup rw_{m+j}$ has full row rank. If $S \cup rw_{m+j}$ has full row rank, then set $S = S \cup rw_{m+j}$. When the rank of $S$ is equal to $n + g$, stop the calculation and return $S$.

The nonzero entries in the last $n + g - m$ rows of $S$ give the optimal solution of SLP model (17). The corresponding roads compose the optimal detection set.

The greedy algorithm can provide the optimal solution to model (17). This can be proven as follows.

**Proposition 9.** Let $rw_{m+k}$ be the row vector with the maximum weight $\omega_k$ among all the row vectors $rw_{m+j}$ in $I_g$ which satisfy that $C_1 \cup rw_{m+j}$ has full row rank. Then, $rw_{m+k}$ must belong to the optimal feasible vector group.

**Proof.** Denote the optimal feasible vector group as $Q = (rw_{m+q_1}, \ldots, rw_{m+q_t}, \ldots, rw_{m+q_{m+n+g}})$, where $q_j \in \{1, 2, \ldots, n+g\}$. Obviously, any row vector in $Q$ will satisfy that $C_1 \cup rw_{m+j}$ has full rank. Assume that $rw_{m+k} \notin Q$. Since $rw_{m+k}$ has the maximal weight among all the row vectors $rw_{m+j}$ in $I_g$ which satisfy that $C_1 \cup rw_{m+j}$ has full row rank, we have $\omega_k \geq \omega_{q_j}$.

Because the vectors in $Q$ can compose a maximal independence row vector group of $C'$ with row vectors in $C_1$, the following equation must hold:

$$
rw_{m+k} = k_1 rw_1 + k_2 rw_2 + \cdots + k_m rw_m + k_{m+1} rw_{m+q_1} + k_{m+2} rw_{m+q_2} + \cdots + k_{m+n+g} rw_{m+q_{m+n+g}} + k_{m+g+1} rw_{m+q_{m+n+g+1}} + \cdots + k_{m+2g} rw_{m+q_{m+2g}}.
$$

where at least one $k_j \neq 0$. Since $C_1 \cup rw_{m+k}$ has full rank, there must be one $k_j \in \{k_{m+1}, k_{m+2}, \ldots, k_{m+g}\}$ which has a nonzero value. Otherwise, if all $k_j \in \{k_{m+1}, k_{m+2}, \ldots, k_{m+g}\}$
Net balancing flow on link 7

(a) Original road network

(b) Road network with dummy road

Figure 4: A simple road network for SLP model.

is zero, then, from (19), \( \text{rw}_{m+k} \) is the linear combination of \( \text{rw}_{j} \) if \( j \in \{1, 2, \ldots, m\} \), which is inconsistent with the fact that \( C_{1} \cup \text{rw}_{m+k} \) has full rank.

Suppose \( k_{m+j} \neq 0 \), \( j \in \{1, 2, \ldots, n + g - m\} \); then, \( C_{1} \cup \text{rw}_{m+k} \) can compose a maximal independence row vector group of \( C' \) with all the vectors in \( Q \) except \( \text{rw}_{m+q_{j}} \). Otherwise, the following equation must hold:

\[
\text{rw}_{m+k} = k_{1}' \text{rw}_{1} + \cdots + k_{m}' \text{rw}_{m} + k_{m+1}' \text{rw}_{m+q_{1}} + \cdots \\
+ k_{m+j-1}' \text{rw}_{m+q_{j-1}} + k_{m+j}' \text{rw}_{m+q_{j}} + \cdots \\
+ k_{m+g}' \text{rw}_{m+q_{m+g}}.
\]

From (19) and (20), we have

\[
0 = (k_{1} - k_{1}') \text{rw}_{1} + \cdots + (k_{m} - k_{m}') \text{rw}_{m} \\
+ (k_{m+1} - k_{m+1}') \text{rw}_{m+q_{1}} + \cdots \\
+ (k_{m+j-1} - k_{m+j-1}') \text{rw}_{m+q_{j-1}} \\
+ (k_{m+j} - k_{m+j}') \text{rw}_{m+q_{j}} + \cdots \\
+ (k_{n+g} - k_{m+g}') \text{rw}_{m+q_{m+g}} + k_{m+j} \text{rw}_{m+q_{j}}, \tag{21}
\]

Since \( k_{m+j} \neq 0 \), (21) is not consistent with the fact that \( Q \cup C_{1} \) is a maximal independence row vector group of \( C' \).

So \( C_{1} \cup \text{rw}_{m+k} \cup (Q - \{\text{rw}_{m+q_{j}}\}) \) is a maximal independence row vector group of \( C' \). Because \( w_{k} \geq w_{q_{j}} \), the total weights of \( \text{rw}_{m+k} \cup (Q - \{\text{rw}_{m+q_{j}}\}) \) must be not smaller than the total weights of \( Q \). If \( Q \) is the optimal feasible vector group, then \( \text{rw}_{m+k} \cup (Q - \{\text{rw}_{m+q_{j}}\}) \) must be the optimal feasible vector group too.

So \( \text{rw}_{m+k} \) must belong to the optimal feasible vector group.

**Proposition 10.** For any \( S \) in the algorithm steps whose rank is smaller than \( n + g \), suppose all the row vectors in \( S \) belong to the optimal feasible vector group. Let \( \text{rw}_{m+k} \notin S \) be the row vector with the maximum weight \( w_{k} \) among all the row vectors \( \text{rw}_{m+j} \in I_{g} - S \) which satisfy that \( S \cup \text{rw}_{m+j} \) has full row rank. \( \text{rw}_{m+k} \) must belong to the optimal feasible vector group.

Proposition 10 can be easily proven if we look at \( S \) here as \( C_{1} \) in Proposition 9. All the proved procedure will be the same with Proposition 9.

Proposition 9 shows that the first row vector which is selected from \( I_{g} \) into \( S \) according to the proposed greedy algorithm must belong to the optimal feasible vector group. Proposition 10 shows that if all the row vectors, which have been selected from \( I_{g} \) into \( S \), compose a subset of the optimal feasible vector group, then the next row vector, which will be selected from \( I_{g} \) into \( S \) according to the proposed greedy algorithm, must belong to the optimal feasible vector group too. So, by the method of mathematical induction, the proposed greedy algorithm is proven to be able to produce the optimal solution to model (17).

5.2.3. Numerical Examples. In this part, a simple road network composed of one exclusive incoming road, two exclusive outgoing roads, and six incoming/outgoing roads is used to examine the greedy algorithm. Another bigger road network with different number of dummy roads is used to show the computation performance.

The simple road network is given in Figure 4.

There is nonzero net balancing flow generated on road 7. The net balancing flow can be transferred to a dummy road 10 as shown in Figure 4(b) with 1 as the turning ratio from road 10 to road 7.

The turning ratios for each road are given by Table 1.

The weight vector is given by \( W = [w_{1}, w_{2}, \ldots, w_{10}] = [4.2, 5.3, 7.8, 7.5, 4.8, 6.5, 3.7, 2.8, 0.1, 5.1]. \)

According to the descending order of weights, the checking order of the roads is \([9, 3, 4, 8, 6, 2, 5, 1, 7, 10]\). By calculating the ranks of matrices, the final optimal detection set is \([9, 6]\), and the maximal weight is 14.5. Actually, after
road 9 is determined to be detected, the flows on roads 3, 4, and 8 are also determined.

A bigger network is given to show the computation performance of the algorithm. The network is shown in Figure 5. This network is composed of 120 directed roads and 25 intersections.

Based on the road network in Figure 5, three scenarios are generated:

(a) No net balancing flow exists in the network.
(b) There are 20 roads whose net balancing flows are nonzero.
(c) There are 40 roads whose net balancing flows are nonzero.

Table 1: Turning ratios.

<table>
<thead>
<tr>
<th>Link number</th>
<th>Downstream links: corresponding turning ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1: 0.7; 6: 0.3</td>
</tr>
<tr>
<td>4</td>
<td>2: 1.0</td>
</tr>
<tr>
<td>5</td>
<td>2: 1.0</td>
</tr>
<tr>
<td>6</td>
<td>2: 1.0</td>
</tr>
<tr>
<td>7</td>
<td>1: 0.4; 6: 0.6</td>
</tr>
<tr>
<td>8</td>
<td>5: 0.2; 7: 0.8</td>
</tr>
<tr>
<td>9</td>
<td>3: 0.5; 4: 0.2; 8: 0.3</td>
</tr>
<tr>
<td>10</td>
<td>7: 1.0</td>
</tr>
</tbody>
</table>

Table 2: CPU time for three scenarios of network A.

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>CPU time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
</tr>
<tr>
<td>(a) No dummy roads</td>
<td>10.546</td>
</tr>
<tr>
<td>(b) 20 dummy roads</td>
<td>16.093</td>
</tr>
<tr>
<td>(c) 40 dummy roads</td>
<td>21.121</td>
</tr>
</tbody>
</table>

We implemented the algorithm for 200 times under each scenario in C# programming language on a Thinkpad (X201s). Every time we generate the turning ratios and weights randomly. The average CPU time for three scenarios is shown in Table 2.

6. Conclusions

The network sensor location problem based on the turning ratios discussed in this paper essentially uses the topology-driven logic to infer the traffic flow over the network. So no assumptions about the route choice are needed. The findings of this paper can be summarized as follows.

First of all, the coefficient matrix of the flow conservation system based on turning ratios is proven to be nonsingular. So the minimal number of the roads required to infer all the traffic flows is determined to be the total number of exclusive incoming roads and the outgoing roads with nonzero net balancing flows.
Second of all, a special sensors location design called "closing ring" is improved in this paper. The analysis shows that if we can detect the traffic flows on all the exclusive incoming roads and all the net balancing flows in the network, then the whole traffic flow vector can be inferred no matter what topology structure the network has. The analysis also shows that we can always find a feasible solution including only real roads for the proposed SLP model, which is useful because the net balancing flow is usually difficult to be detected.

And finally, a weighted SLP model to find the optimal detecting set considering the roads importance and cost of sensors is formulated. Greedy algorithm is designed to solve the model, which is proven to be able to provide the optimal solution. Two numerical examples are given to test the algorithm. The computation performance is reported.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The work of this paper is supported by the National Natural Science Foundation of China (Grant no. 51208379).

References
