Article

Sensor Deployment Strategy and Traffic Demand Estimation with Multisource Data

Hui Chen [1,2], Zhaoming Chu [3,*] and Chao Sun [4]

1   College of Automobile and Traffic Engineering, Nanjing Forestry University, Nanjing 210037, China; 13621092886@163.com
2   Department of Infrastructure Development, National Development and Reform Commission, Beijing 100045, China
3   Research Institute for Road Safety of MPS, Beijing 100062, China
4   School of Automotive and Traffic Engineering, Jiangsu University, Zhenjiang 212013, China; chaosun@ujs.edu.cn
*   Correspondence: chuzhaoming@126.com; Tel.: +86-137-1790-6369

Abstract: Since traffic origin-destination (OD) demand is a fundamental input parameter of urban road network planning and traffic management, multisource data are adopted to study methods of integrated sensor deployment and traffic demand estimation. A sensor deployment model is built to determine the optimal quantity and locations of sensors based on the principle of maximum link and route flow coverage information. Minimum variance weighted average technology is used to fuse the observed multisource data from the deployed sensors. Then, the bilevel maximum likelihood traffic demand estimation model is presented, where the upper-level model uses the method of maximum likelihood to estimate the traffic demand, and the lower-level model adopts the stochastic user equilibrium (SUE) to derive the route choice proportion. The sequential identification of sensors and iterative algorithms are designed to solve the sensor deployment and maximum likelihood traffic demand estimation models, respectively. Numerical examples demonstrate that the proposed sensor deployment model can be used to determine the optimal scheme of refitting sensors. The values estimated by the multisource data fusion-based traffic demand estimation model are close to the real traffic demands, and the iterative algorithm can achieve an accuracy of 10^{-3} in 20 s. This research has significantly promoted the effects of applying multisource data to traffic demand estimation problems.

Keywords: traffic demand estimation; multisource data; sensor deployment; sequential identification; iterative algorithm

1. Introduction

Traffic demand describes the spatial distribution characteristics of travel of passengers and goods, which reflects the basic information of motor vehicles' movement between traffic zones. Traffic demand is the most basic input parameter in urban long-term transportation planning and short-term traffic management and control. Urban traffic route guidance, speed limits, congestion control and other measures require an accurate origin-destination (OD) demand matrix. Traffic demand not only provides a decision-making basis for the balance of supply and demand of urban road networks and government construction investment, but also provides data support for the development of urban motorization. Only by accurately estimating the OD demand can we grasp the traffic characteristics of the current road network and provide more targeted and effective strategies to ease traffic congestion in the processes of urban planning and transportation planning.

With the development of urban traffic information acquisition technology, we can collect observation data from a variety of traffic detectors. Detector types include loop-coil detection, ultrasonic detection, and image recognition detection. The detection data types

mainly include road flow, speed, density, travel time, vehicle trajectory, etc. This paper aims to estimate traffic demand using these multisource data.

In recent years, scholars have used a variety of theories to estimate traffic demand, and the main research methods include those listed below. (1) The generalized least square method [1,2]. In this method, the least square sum of the errors between the observed variables and the estimated variables is found by modifying the OD demand matrix to be estimated. Because the mathematical formula of this method is simple and easy to understand, it is widely used in OD estimation. (2) The OD estimation method based on information entropy [3]. The objective function of the least square method is the minimum distance between the observed variable and the estimated variable, while the objective function of the information entropy model is the maximum entropy or the minimum information to estimate the OD demand. (3) The maximum likelihood estimation [4,5]. The objective of this method is to maximize the likelihood function based on the observed flow and the historical OD matrix to estimate the OD matrix. (4) The Bayesian estimation method [6]. This method is the same as the classic statistical method, assuming that the observed variables are random variables, and the Bayesian method also considers that the parameters to be estimated are random variables. The principle is to give the prior distribution of all the variables to be estimated and then derive the posterior distribution of all the variables to be estimated according to the collected sample information to give the estimated value and the corresponding confidence interval of the variables to be estimated. The derivation of a posteriori distribution often requires some complex integral calculus. The Monte Carlo simulation algorithm is typically used to sample and solve the estimation and confidence interval of a posteriori distribution. For other traffic network OD estimation models with new observation variables, we can refer to the OD estimation model based on Smartcard data [7], cellular data [8], license plate recognition data [9] and the OD estimation model based on mobile location data [10].

As a subproblem of traffic demand, the network detector placement problem can be used to identify the location of detectors to determine the unique unobserved road flow or to optimize the quality of traffic demand to be estimated. To obtain the only unobserved link flow, He [11] used graph theory to solve the optimal detector layout strategy. Based on algebra method, some scholars build the corresponding network detector layout model [12] according to the link, path and node correlation matrix. To optimize the quality of the parameters to be estimated, Simonelli et al. [13] used the trace of the covariance of the posterior OD matrix to determine the optimal detector placement strategy based on Bayesian statistics. Yang and Zhou [14] further proposed the maximum coverage criterion of path flow. Salari et al. [15] analyzed the effect of sensor failure on the detector layout strategy.

Although different methods have been used to study the detector layout strategy and traffic demand estimation, the following deficiencies exist: (1) research on traffic demand estimation using multisource data fusion technology has seldom been used in the existing literature. (2) Research on the layout of traffic detectors does not consider the coverage information of road sections and path flows. (3) The lack of an analytical relationship between traffic demand and traffic flow from the perspective of traffic network theory and the relationship is helpful to establish the maximum likelihood traffic demand estimation model.

To address these deficiencies, this paper proposes an integrated model of detectors layout and maximum likelihood traffic demand estimation. The former calculates the optimal number and location of the detector layout by maximizing the road section and path coverage information and uses minimum variance weighted average technology to fuse the detected multisource data to use the maximum likelihood method to estimate the traffic demand. The successive detector identification algorithm and iterative algorithm are designed to solve the detector layout and traffic demand estimation model, and the Nguyen Dupuis and Sioux Falls networks are used to test the model and algorithm [16,17]. The contributions of this paper mainly contains: (1) proposed an integrated model of

detectors layout and maximum likelihood traffic demand estimation. (2) designed the successive detector identification algorithm and iterative algorithm.

2. Detector Layout and Traffic Demand Estimation

2.1. Detector Layout Model

Traffic demand estimation can be regarded as the following parameter estimation problem in mathematics: some samples (such as road flow, speed, density, travel time, etc.) are observed to estimate parameters in the overall distribution (traffic demand). Under the current budget, road samples are observed to estimate the most reliable or most practical traffic demand, which is the premise scientific problem to be solved in this paper.

We determine the optimal detector layout quantity and location by studying the detector layout model, and the existing detector layout models unilaterally maximize the link coverage information or path coverage information. Under the given budget, this paper maximizes the coverage information of links and paths at the same time:

$$max \left[\alpha \sum_{a \in A} \widetilde{v}_a z_a + (1-\alpha) \sum_{\omega \in W} \sum_{k \in K^\omega} \widetilde{f}_k^\omega \Psi_k^\omega \right] \tag{1}$$

subject to

$$b_s s \leq b_{max} \tag{2}$$

$$\sum_{a \in A} z_a = s \tag{3}$$

$$\sum_{a \in A} \delta_{a,k}^\omega z_a \geq \Psi_k^\omega, \forall \omega \in W, k \in K^\omega \tag{4}$$

$$z_a \in \{0,1\}, \forall a \in A \tag{5}$$

$$\Psi_k^\omega \in \{0,1\}, \forall \omega \in W, k \in K^\omega \tag{6}$$

where W is a collection of all OD pairs in the transportation network and ω is one of those OD pairs. K^ω is the collection of all paths between OD pair ω, $k \in K^\omega$; α is the weight coefficient; $\widetilde{v}_a$ represents the prior flow on segment a; $\widetilde{f}_k^\omega$ is the prior flow on path k between OD pair ω; Ψ_k^ω is the coverage factor on path k between OD pair ω. The parameter is 1 if there is a road section with detectors in path k; otherwise, it is 0; s is the number of detectors; b_s is the cost of installing detectors for each road section; b_{max} is the general budget; z_a is the coverage factor on road section a, the parameter is 1 if there is the road section a with detectors, otherwise it is 0; A is the Section set, $a \in A$; $\delta_{a,k}^\omega$ is the Path link association parameters, between OD pair ω, the parameter is 1 if the road section a is on path k; otherwise, it is 0.

Objective function (1) maximizes the coverage information of road sections and paths, where $\sum_{a \in A} \widetilde{v}_a z_a$ is the sum of the traffic volume of the road sections where the detectors are deployed; that is, the detectors will be preferentially deployed on the road sections with high traffic flow of the prior road sections. $\sum_{\omega \in W} \sum_{k \in K^\omega} \widetilde{f}_k^\omega \Psi_k^\omega$ is the sum of the traffic volume of the path where the detectors are deployed. The difference between the two is that the objective function calculated by the former is larger because the former will repeatedly stack the path flow, while the latter considers the repeatability between paths. Equation (2) is the cost constraint of the detector layout; Equation (3) is the constraint of the total number of road detectors; Equation (4) is the relationship between the road section and path coverage information; and Equations (5) and (6) are the road section coverage factor and path coverage factor, respectively.

2.2. Multisource Observation Variable Fusion

According to the detector layout model, the number and location of detectors are calculated, and the traffic flow, speed, density and travel time are observed. Because the

dimensions of the observed variables are not consistent, traffic flow theory is used to transform the dimensions of the observed variables to unify the observed multisource road information. In this paper, the green shields traffic flow speed density relation function is used to unify the multisource observation variables into link flow:

$$v_a = \frac{l_a}{t_a(v_a)}\left[M_{j,a}\left(1 - \frac{l_a}{t_a(v_a)S_{f,a}}\right)\right], \forall a \in A \tag{7}$$

where v_a is the traffic volume on segment a, t_a is the estimated travel time on segment a, l_a is the length on segment a, $M_{j,a}$ is the crowding density on segment a, and $S_{f,a}$ is the maximum speed on segment a. Road travel time can be derived from road length and vehicle speed.

After the dimension is unified as the traffic volume of the road section, the method of minimum variance weighted average is used to fuse the multisource road flow to comprehensively reflect the traffic status of the road. The mathematical expression of the minimum variance weighted average method is as follows:

$$\widehat{V} = \sum_{i \in I} e_i v_i = e^T v \tag{8}$$

The flow weight coefficient e_i of each type of observation section is as follows:

$$e_i = \frac{1}{\sigma_i^2 \cdot \sum_{i \in I} \frac{1}{\sigma_i^2}}, \forall i \in I \tag{9}$$

where i is the link flow type, such as speed, density, time, etc., and $i \in I$; v_i is the i-th type of road section flow; and v is its vector form, i.e., $v = \{\cdots, v_i, \cdots\}$; e_i is the i-th type of weight coefficient, and e is its vector form, i.e., $e = \{\cdots, e_i, \cdots\}$; from Formula (9), $\sum_{i \in I} e_i = 1$; σ_i^2 is the variance of i-th type observed flow.

2.3. Maximum Likelihood Traffic Demand Estimation Model

Because the arrival of vehicles in unit time is random, this paper uses a Poisson distribution to describe the random traffic demand $q^\omega \sim \mathrm{Pois}(d^\omega)$. If the traffic volume is not very low, the traffic demand is approximately in the form of a multivariate normal distribution as follows:

$$q \sim N(d, \Lambda) \tag{10}$$

where d is the mean vector form of traffic demand, i.e., $d = \{\cdots, d^\omega, \cdots\}$. This paper assumes that the traffic demand is independent of each other, so the covariance matrix of traffic demand is $\Lambda = \mathrm{diag}(d)$.

In the stochastic user equilibrium model, the relationship between path flow and OD demand is as follows:

$$f_k^\omega = q^\omega \cdot p_k^\omega(c^\omega) = q^\omega \cdot \mathrm{Pr}(c_k^\omega \le c_l^\omega, \forall l \in K^\omega, l \ne k | c^\omega), \forall \omega \in W, k \in K^\omega \tag{11}$$

where f_k^ω is the flow on path k between OD pair ω; $\mathrm{Pr}(\cdot)$ is the probability operator; c^ω is the vector form of path travel time c_k^ω, i.e., $c^\omega = \{\cdots, c_k^\omega, \cdots\}$; where p_k^ω is the route selection probability of travel users on path k between OD pair ω and it can be obtained by calculating the following logit formula:

$$p_k^\omega = \frac{\exp(-\theta c_k^\omega)}{\sum_{k \in K^\omega} \exp(-\theta c_k^\omega)}, \forall \omega \in W, k \in K^\omega \tag{12}$$

where θ is the discrete parameter used to measure the degree of perception error of road travelers; where c_k^ω is the estimated route travel time on path k between OD pair ω. Therefore, the path flow also follows the form of a multivariate normal distribution is as follows:

$$f \sim N\left(pd, p\Lambda p^T\right) \tag{13}$$

where p is the mean vector form of path selection probability, i.e., $p = \left\{\cdots, p_k^\omega, \cdots\right\}$. Because the road flow is the sum of the path flow, the central limit theorem shows that the road flow also obeys the form of a multivariate normal distribution is as follows:

$$V \sim N\left(\Delta pd, \Delta p\Lambda p^T\Delta^T\right) \tag{14}$$

where V is the mean vector form of flow of all sections, i.e., $V = \left\{\cdots, v_a, \cdots\right\}$, v_a is the traffic flow on the road section; A is the section set; $\Delta = \left(\delta_{a,k}^\omega\right)$ is the path link incidence matrix.

According to the fusion of the observed link flow (Equation (8)) and the analytical relationship between link flow and traffic demand (Equation (14)), the likelihood function of traffic demand is established as follows:

$$L(d) = g(V|d) \tag{15}$$

where $g(\cdot)$ is the probability density equation. It can be seen from the front that the road section flow V obeys the multivariate normal distribution, and its log likelihood function is as follows:

$$lnL(d) = -\frac{1}{2}ln\left(\left|\Delta p\Lambda p^T\Delta^T\right|\right) - \frac{1}{2}\left(\widehat{V} - \Delta pd\right)^T\left(\Delta p\Lambda p^T\Delta^T\right)^{-1}\left(\widehat{V} - \Delta pd\right) \tag{16}$$

where $\widehat{V}$ is the fused observed link flow.

Based on the fusion technology, i.e., the link flow obtained from Equations (8) and (9), the maximum likelihood traffic demand estimation model is established by using the bi-level programming theory, the upper model uses the maximum likelihood method to solve the traffic demand, and the lower model uses the SUE model to solve the path selection probability and the upper level is as follows:

upper level:

$$max[lnL(d)] \tag{17}$$

lower level: The SUE model (Equation (11)) is used to allocate the traffic demand to obtain the route selection probability.

3. Algorithm Design

To identify the number and location of detectors, a successive identification detector algorithm is designed to solve the detector layout model; that is, one optimal detector location is identified each time and added to the detector set, and the path coverage information is updated. This method is simple and easy to apply to practical engineering. When solving the maximum likelihood traffic demand estimation model, the iterative algorithm framework is used to calculate the upper and lower models repeatedly. The upper maximum likelihood model is solved by the steepest descent method, and the lower SUE model is solved by a successive average algorithm. The algorithm is as follows (Algorithm 1):

Algorithm 1: Successive recognition detector algorithm

step 1.1	Initialization: set the number of initial detectors $s = 0$; weight coefficient α; prior traffic demand $\boldsymbol{D}$; initial detector placement strategy $\boldsymbol{Z} = \varnothing$; initial path set $\boldsymbol{Y} = \varnothing$.
step 1.2	Solving SUE model with successive average algorithm, obtain road flow v_a, path flow f_k^ω and path link association parameters $\delta_{a,k}^\omega$; according to Equation (2), calculate the maximum number of detectors s_{max} in the road section.
step 1.3	Link model coefficient: for each road section, the detector layout coefficient $\alpha \cdot \widetilde{v}_a$ based on road section flow is calculated.
step 1.4	Path model coefficient: in the whole road network, the path flow in the path set $\boldsymbol{Y}$ is eliminated to obtain the new link flow $\widetilde{v}_{route,a}$, then the detector layout coefficient $(1 - \alpha) \cdot \widetilde{v}_{route,a}$ based on path flow is calculated for each road section.
step 1.5	Identifying a new detector position $Z^{(s)}$ (except $\boldsymbol{Z}$) makes the objective function maximum, that is $max_{Z^{(s)} \in A} \left[\alpha \cdot \widetilde{v}_{Z^{(s)}} + (1 - \alpha) \cdot \widetilde{v}_{route,Z^{(s)}} \right]$, and add it to the detector layout strategy $\boldsymbol{Z}$. In addition, add all paths through the section to the path collection $\boldsymbol{Y}$.
step 1.6	Convergence test: order $s = s + 1$, if $s \geq s_{max}$, stop the calculation, $\boldsymbol{Z}$ is the optimal detector placement strategy; otherwise, go to step 1.4.

According to the number and location of detectors obtained by Algorithm 1, observe traffic flow, speed, density, travel time, etc. on these sections. Then, Formulas (8) and (9) are used to fuse these observation variables, and traffic demand is estimated by Algorithm 2.

Algorithm 2: Iterative algorithm for traffic demand estimation model

step 2.1	Initialization: set the number of iteration steps $n = 0$; convergence accuracy ε; mean $\boldsymbol{d}^{(0)}$ of traffic demand, so the covariance matrix of traffic demand is $\Lambda = diag\left(\boldsymbol{d}^{(0)} \right)$, at the same time, the initial traffic demand is set as $\boldsymbol{d}^{(0)}$; link flow $\hat{\boldsymbol{V}}$ after fusion.
step 2.2	Solving the underlying model: Using successive average algorithm to solve SUE model, i.e., distribution demand $\boldsymbol{d}^{(n)}$, then the path selection probability $\boldsymbol{p}^{(n)}$ is obtained.
step 2.3	Solving the upper model: probability of path selection in $\boldsymbol{p}^{(n)}$, using the steepest descent method to solve the auxiliary OD demand $\overline{\boldsymbol{d}}^{(n)}$.
step 2.4	Update traffic requirements: order $\boldsymbol{d}^{(n+1)} = \boldsymbol{d}^{(n)} + \left(\frac{1}{n} \right) \cdot (\overline{\boldsymbol{d}}^{(n)} - \boldsymbol{d}^{(n)})$.
step 2.5	Convergence test: if $\|\boldsymbol{d}^{(n+1)} - \boldsymbol{d}^{(n)}\| / \|\boldsymbol{d}^{(n)}\| \leq \varepsilon$, stop the calculation, $\boldsymbol{d}^{(n+1)}$ is the optimal traffic demand; otherwise, order $n = n + 1$, go to step 2.2.

The designed algorithms are programmed in MATLAB R2016a and tested on the notebook computer with Intel(R) Core(TM) i7-5600U CPU 2.60 GHz, 8 GB memory.

4. Example Analysis

4.1. Nguyen-Dupuis Network

Nguyen-Dupuis Network [16] is composed of 13 nodes, 38 road sections and 18 OD pairs. The topology, road section attributes and traffic demand are shown in Figure 1, Tables 1 and 2, respectively. The number near the line in Figure 1 represents road order. The section impedance is represented by the section characteristic function $t_a(x_a) = t_a^0 \left[1 + \beta \left(\frac{x_a}{C_a} \right)^n \right]$ of the Federal Highway Administration (BRP), where t_a, t_a^0, x_a and C_a represent the actual travel time, free travel time, flow and capacity of link a, respectively, and $\beta = 0.15$ and $n = 4$ are the determining parameters.
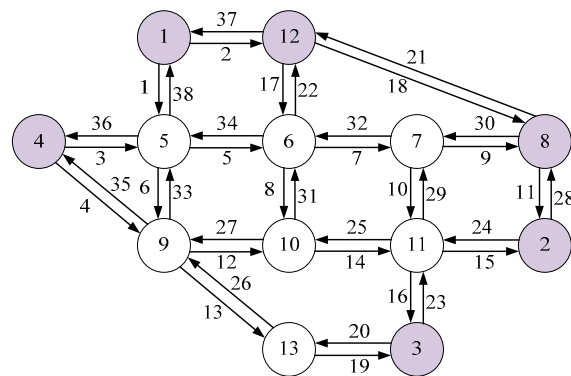
Figure 1. Nguyen-Dupuis network topology.

Table 1. Nguyen-Dupuis network link properties.

Link	t_a^0	C_a	Link	t_a^0	C_a	Link	t_a^0	C_a
1	7	700	14	4	700	27	4	280
2	9	560	15	9	280	28	9	700
3	12	560	16	8	560	29	4	700
4	5	375	17	7	140	30	9	700
5	12	420	18	14	560	31	4	280
6	9	420	19	11	560	32	5	700
7	5	700	20	11	560	33	9	420
8	4	280	21	14	560	34	12	420
9	9	700	22	6	140	35	5	375
10	4	700	23	9	560	36	12	560
11	9	700	24	9	280	37	9	560
12	5	280	25	4	700	38	7	700
13	9	280	26	9	280			

Table 2. Nguyen-Dupuis network OD demand.

O	D	Demand	O	D	Demand
1	2	210.00	2	1	210.00
1	3	430.00	2	4	320.00
1	8	320.00	2	12	50.00
4	2	320.00	3	1	430.00
4	3	110.00	3	4	110.00
4	8	210.00	3	12	40.00
12	2	50.00	8	1	320.00
12	3	40.00	8	4	210.00
12	8	60.00	8	12	60.00

First, the detector layout model is tested. The prior OD demand is shown in Table 2, the weight coefficient is set to $\alpha = 0.5$, and the maximum number of detectors (section speed) of road section $s_{max} = 6$. Figure 2 depicts the changes in the objective function values of the three models (the objective function only considers the road section coverage information: $\sum_{a \in A} \widetilde{v}_a z_a$, a detector layout model in this paper, and the objective function only considers path coverage information: $\sum_{\omega \in W} \sum_{k \in K^\omega} \widetilde{f}_k^\omega \Psi_k^\omega$) after the detector layout position is identified one by one. The target function values of the three models increase with the increase in the number of detectors, which means that more observation information can obtain a higher coverage information detector deployment scheme. Compared with the traditional detector layout model, the detector layout model proposed in this paper can obtain a more reliable detector layout scheme by using the coverage information of both road sections and paths.
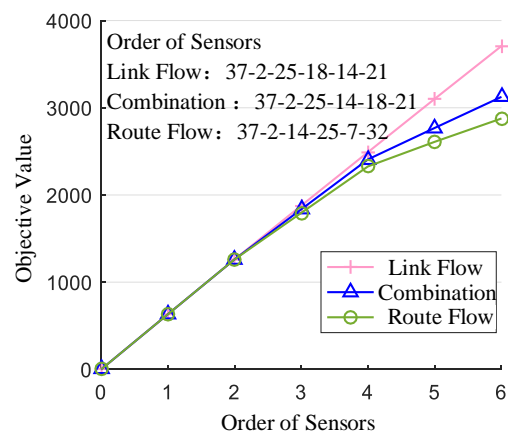
Figure 2. Under different models, the updated objective function values are identified one by one with the detector location.

To illustrate the robustness of the maximum likelihood traffic demand estimation model, Figure 3 analyses the sensitivity of traffic demand. The root mean square error (*RMSE*) of traffic demand is defined as follows:

$$RMSE = \frac{\|\boldsymbol{d}^{tru} - \boldsymbol{d}^{est}\|}{\sqrt{|\omega|}} \tag{18}$$

where $\boldsymbol{d}^{tru}$, $\boldsymbol{d}^{est}$ and $|\omega|$ represent the accurate traffic demand, the estimated traffic demand and the potential of traffic demand respectively. The average value of traffic demand is shown in Table 2. According to the detector location calculated by the previous detector layout model, the flow, speed, density, travel time, etc. on observation Sections 37, 2, 25, 14, 18 and 21 are obtained from the average value of traffic demand allocated by the SUE model. These observed traffic variables are further fused according to Equation (8). To illustrate the effectiveness of the model, $\boldsymbol{d}$ is also used as the "accurate traffic demand" of the test network. It can be seen in the figure that the *RMSE* of traffic demand reaches the minimum value of 0 when the observed link flow and speed disturbance are both 0%, which indicates that the proposed maximum likelihood traffic demand estimation model can estimate the traffic demand close to the actual value. *RMSE* increases with the increase in link flow and speed disturbance, which means that the disturbance of input parameters (link flow and speed) has a significant impact on the estimation of traffic demand.
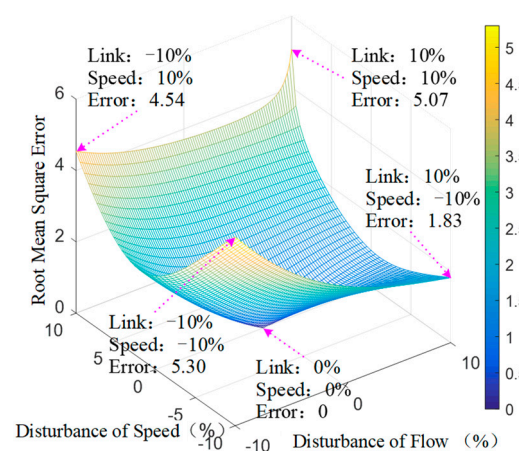

Figure 3. Influence of root mean square error of demand estimation on observed road flow and speed.

4.2. Sioux Falls Network

Sioux Falls network [17] is used to illustrate the convergence characteristics of the design algorithm. The network consists of 24 nodes, 76 road sections and 550 OD pairs (as shown in Figure 4). The BPR characteristic function ($\beta = 0.15$, $n = 4$) is also used to calculate the section impedance, and the link characteristics and mean OD demand reference to [17].
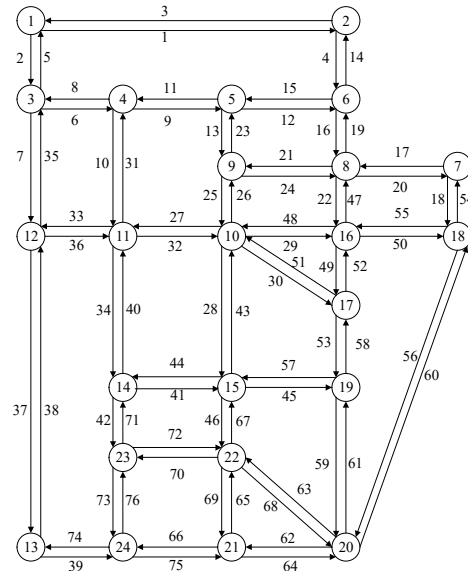

Figure 4. Sioux Falls network topology.

In the actual road network, some road sections have been equipped with detectors. The optimal layout strategy of modified and newly installed detectors is tested in Figure 5. Detectors have been installed on Sections 11, 34, 43, 58 and 72 with a total budget of 400. The costs of retrofitting and installing new detectors are 50 and 100, respectively. Figure 5 shows that the objective function value increases with the increase in the number of modified or new detectors. Compared with the existing detectors, the modified or new detectors produce higher objective function values, which means that the modified or new detectors can obtain more complete coverage information to calculate more reliable traffic demand. Figure 5 also shows that in the optimal detector layout state, two detectors need to be modified, three detectors need to be added and three detectors need to be retained. Therefore, the detector layout model can be used to formulate the layout strategy of refitting and adding detectors with maximum coverage information under a given budget.
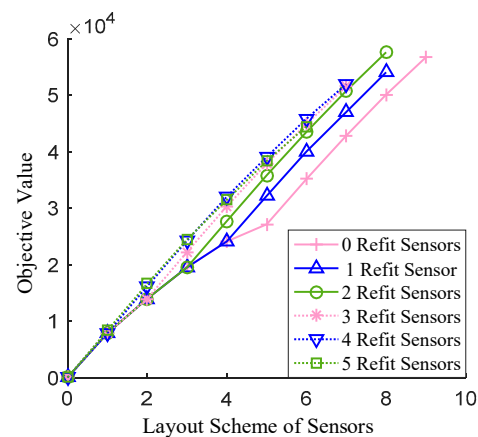

Figure 5. Objective function value identified one by one with the location of the detector in different numbers of refitted detectors.

Figures 6 and 7 further describe the convergence characteristics of the design iterative algorithm. The iterative algorithm can achieve an accuracy of 10^{-3} in 20 s. Similarly, the successive average algorithm for solving the lower SUE model, the one-dimensional search algorithm and the steepest descent algorithm for solving the upper maximum likelihood estimation can quickly converge to the equilibrium solution. One-dimensional search is the step search method in the steepest descent method. The solution of this process has an oscillation phenomenon, but the convergence error shows a downward trend as a whole, and other algorithms can converge steadily. Figure 7 tests the evolution process of traffic demand. The initial iteration process of traffic demand changes greatly. When running for approximately 150 s, the traffic demand can converge to the accurate value.
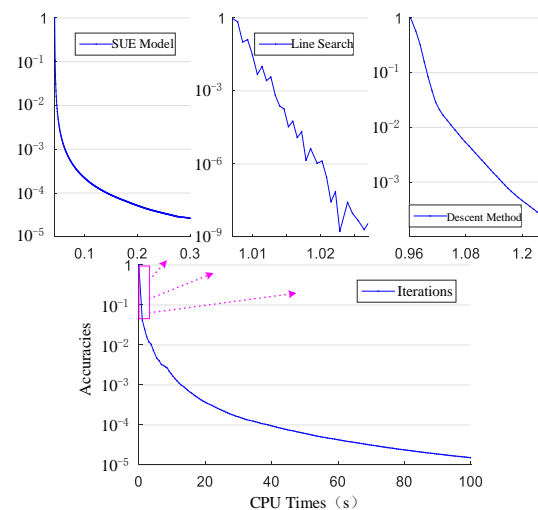

Figure 6. Convergence curves of the iterative algorithm and successive average algorithm, one-dimensional search and steepest descent algorithm in the first iteration process.
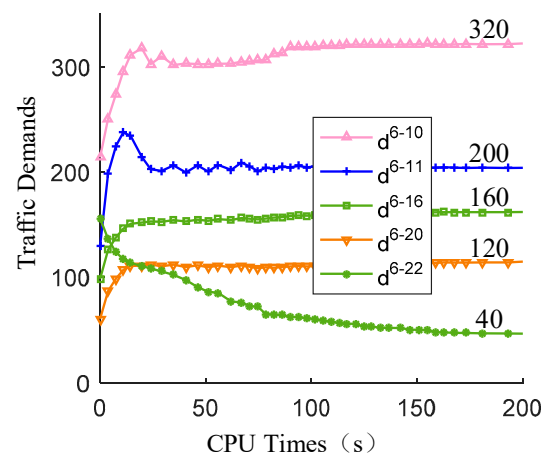

Figure 7. Estimated traffic demand evolution curve.

5. Conclusions

To obtain the basic data of urban road network planning and traffic control, a detector layout model and a maximum likelihood traffic demand estimation model are established by using the multisource data observed by modern information technology. The former provides the location and quantity of data monitoring for the latter, and the minimum variance weighted average technology is used to fuse the multisource data. Furthermore, the bilevel programming method is used to estimate the traffic demand. We design a successive detector recognition algorithm and an iterative algorithm to solve the two models.

We use the Nguyen Dupuis and Sioux Falls networks to test the performance of the proposed model and algorithm. The results show that the detector placement scheme

considering both the coverage information of road section and path is more reliable, and the designed detector layout model can determine the optimal layout strategy for newly installed and refitted detectors. The disturbance of input parameters has a significant impact on traffic demand estimation, and the designed algorithm can achieve an accuracy of 10^{-3} in 20 s; the traffic demand estimation method based on multisource data fusion fully mines the observed data information to estimate the traffic demand close to the actual value. In a follow-up study, we will estimate the dynamic traffic demand in the congestion network and further use the estimated traffic demand to make proper traffic management policies. And apply the proposed model in the real urban traffic network.

Author Contributions: Conceptualization, H.C. and C.S.; methodology, H.C. and C.S.; software, H.C.; validation, H.C. and C.S.; formal analysis, H.C.; investigation, Z.C.; resources, Z.C.; data curation, H.C.; writing—original draft preparation, H.C.; writing—review and editing, Z.C.; visualization, C.S.; supervision, Z.C.; project administration, Z.C.; funding acquisition, C.S. and Z.C. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cao, Y.; Tang, K.; Sun, J.; Ji, Y. Day-to-day dynamic origin-destination flow estimation using connected vehicle trajectories and automatic vehicle identification data. *Transp. Res. Part C Emerg. Technol.* **2021**, *129*, 103241. [CrossRef]
2. Caggiani, L.; Ottomanelli, M.; Sassanelli, D. A fixed point approach to origin-destination matrices estimation using uncertain data and fuzzy programming on congested networks. *Transp. Res. Part C Emerg. Technol.* **2013**, *28*, 130–141. [CrossRef]
3. William, H.K.; Shao, H.; Cao, S.; Yang, H. Origin-destination demand estimation models. *Int. Encycl. Transp.* **2021**, *4*, 515–518.
4. Spiess, H. A maximum likelihood model for estimating origin-destination matrices. *Transp. Res. Part B Methodol.* **1987**, *21*, 395–412. [CrossRef]
5. Mohanty, S.; Pozdnukhov, A. Dynamic origin-destination demand estimation from link counts, cellular data and travel time data. *Transp. Res. Procedia* **2020**, *48*, 1722–1739. [CrossRef]
6. Mo, B.; Li, R.; Dai, J. Estimating dynamic origin–destination demand: A hybrid framework using license plate recognition data. *Comput.-Aided Civ. Infrastruct. Eng.* **2020**, *35*, 734–752. [CrossRef]
7. Parry, M.; Hazelton, M.L. Estimation of origin-destination matrices from link counts and sporadic routing data. *Transp. Res. Part B Methodol.* **2012**, *46*, 175–188. [CrossRef]
8. Zhu, S.; Zheng, H.; Peeta, S.; Guo, Y.; Cheng, L.; Sun, W. Optimal heterogeneous sensor deployment strategy for dynamic origin-destination demand estimation. *Transp. Res. Rec. J. Transp. Res. Board* **2016**, *2567*, 18–27. [CrossRef]
9. Hussain, E.; Bhaskar, A.; Chung, E. Transit OD matrix estimation using smartcard data: Recent developments and future research challenges. *Transp. Res. Part C Emerg. Technol.* **2021**, *125*, 103044. [CrossRef]
10. Ge, Q.; Fukuda, D. Updating origin-destination matrices with aggregated data of GPS traces. *Transp. Res. Part C Emerg. Technol.* **2016**, *69*, 291–312. [CrossRef]
11. He, S. A graphical approach to identify sensor locations for link flow inference. *Transp. Res. Part B Methodol.* **2013**, *51*, 65–76. [CrossRef]
12. Xu, X.; Lo, H.K.; Chen, A.; Castillo, E. Robust network sensor location for complete link flow observability under uncertainty. *Transp. Res. Part B Methodol.* **2016**, *8*, 1–20. [CrossRef]
13. Simonelli, F.; Marzano, V.; Papola, A.; Vitiello, I. A network sensor location procedure accounting for O-D matrix estimate variability. *Transp. Res. Part B Methodol.* **2012**, *46*, 1624–1638. [CrossRef]
14. Yang, H.; Zhou, J. Optimal traffic counting locations for origin-destination matrix estimation. *Transp. Res. Part B Methodol.* **1998**, *32*, 109–126. [CrossRef]
15. Salari, M.; Kattan, L.; Lam, W.H.K.; Esfeh, M.A.; Fu, H. Modeling the effect of sensor failure on the location of counting sensors for origin-destination (OD) estimation. *Transp. Res. Part C Emerg. Technol.* **2021**, *132*, 103367. [CrossRef]
16. Nguyen, S.; Dupuis, C. An efficient method for computing traffic equilibria in networks with asymmetric transportation costs. *Transp. Sci.* **1984**, *18*, 185–202. [CrossRef]
17. Leblanc, L.J. Mathematical Programming Algorithms for Large Scale Network Equilibrium and Network Design Problems. Ph.D. Thesis, Northwestern University, Evanston, IL, USA, 1973.