THE APPLICATION OF OD FLOW ESTIMATION MODEL FROM OBSERVED LINK TRAFFIC FLOW:EMPIRICAL EVIDENCE IN TAINAN CITY

Hao-Ching Hsia, College of planning and design, National Cheng Kung University, Taiwan, spacemake@gmail.com

Kuang-Yih Yeh, College of planning and design, National Cheng Kung University, Taiwan

Fumitaka Kurauchi, Faculty of Engineering, Gifu University, Japan

Nobuhiro Uno, Graduate School of Management, Kyoto University, Japan

Yan Li, College of Asia Pacific Studies, Ritsumeikan Asia Pacific University, Japan

Hiroshi Tsukaguchi, Faculty of Engineering, Ritsumeikan University, Japan

Yasunori Iida, Institute of System Science Research, Japan

ABSTRACT

This study aims at establishing an origin-destination (OD) flow estimation model. The model originates from observed link traffic flow for solving problems traditional model faced and providing a practical tool for transportation planners. The advantages of the model developed in this study include: 1) the integration of dummy traffic zones and actual road network is sufficient; 2) all the inner-inner, inner-outer, outer-inner and outer-outer OD traffic volume can be calculated at the same time; 3) the OD traffic volume is described as a function of the trip production only, and 4) the difference between estimated and measured link traffic volumes is minimized. In order to confirm the practicability of this concept, link traffic flow data on the actual road network was collected in Tainan city. A numerical experiment was conducted to

examine the reliability and validity of the proposed model from the viewpoint of accuracy of estimated OD traffic volumes.
Keywords: origin destination flow estimation model, observed link traffic flow, trip production

INTRODUCTION

Understanding traffic demand distribution or original-destination (OD) matrix is very essential and important for every kind of transport planning. Various tools have obtained to collect more details and up-to-date information by ITS technologies. Still, the questionnaire-based survey has been adopted to estimate the traffic demand distribution. However, by the increasing interest in the privacy issues, the response rate for the questionnaire survey decreases, resulting in the lower accuracy of the estimated result from the survey.

On the other hand, the estimation method of OD matrices has been studied since 60s, and large variations of models have been proposed. Many OD matrices estimation models have been developed, but some of them adopt strict or unrealistic assumptions that make them difficult to apply to a real-world network. In this study, on behalf of looking at state-of-art technologies of the estimation method, we focus on the capability of model implementation to the real network.

As the explanation above, we are not interested in proposing a novel model of OD matrices in this study but validating whether the existing model can be applicable for the practical use. For this reason, we are not looking at complex analytical models such as ‘dynamic OD estimation (Sherali et al., 1997; Sherali et al., 2001; Willumsen, 1984; Oneyama and Kuwahara, 1997; Cascetta et al., 1993 and Cremer and Keller, 1987)’, ‘OD estimation with stochastic link traffic volumes (Davis and Nihan, 1991)’, ‘bi-level optimization formulation considering the congestion effect (Yang et al., 1992)’, or ‘sequential linear programming approach’ (Sherali et al. 2003).

This paper consists of 5 chapters. Chapter 1 described the objective of this study. Chapter 2 explains the formulation of the OD estimation model. In Chapter 3, the developed model is verified by using the simple hypothetical network. Key factors influencing the estimation results are identified by the sensitivity analysis. Further in Chapter 4, the proposed model is further applied to the Tainan network. Chapter 5 then concludes the paper.

FORMULATION

Notation

n_i: Number of inner origins

n_o: Number of outer origins

m_i: Number of inner destinations

m_o: Number of outer destinations

a: Number of links

w_{ii}: Number of inner-inner OD pairs (OD pairs originated from inner node destined also to the inner node.)

w_{io}: Number of inner-outer OD pairs

w_{oi}: Number of outer-inner OD pairs

w_{oo}: Number of outer-outer OD pairs

τ: Rate of the traffic generated within the area destined within the area ($0 \leq \tau \leq 1$)

λ: Rate of the traffic entered from outside destined within the area ($0 \leq \lambda \leq 1$)

g: A vector of generated traffic volumes at the inner origins with the size of n_i

s: A vector of generated traffic volumes at the outer origins with the size of n_o

D_{ii}: A OD-origin incidence matrix for inner-inner OD pairs with the size of w_{ii} x n_i

D_{io}: A OD-origin incidence matrix for inner-outer OD pairs with the size of w_{io} x n_i

D_{oi}: A OD-origin incidence matrix for outer-inner OD pairs with the size of w_{oi} x n_o

D_{oo}: A OD-destination incidence matrix for outer-outer OD pairs with the size of w_{oo} x n_o

Q_{ii}: A OD-destination incidence matrix for inner-inner OD pairs with the size of w_{ii} x m_i

Q_{io}: A OD-destination incidence matrix for inner-outer OD pairs with the size of w_{io} x m_i

Q_{oi}: A OD-destination incidence matrix for outer-inner OD pairs with the size of w_{oi} x m_o

Q_{oo}: A OD-destination incidence matrix for outer-outer OD pairs with the size of w_{oo} x m_o

m: A vector of destination choice probability for inner-inner centroids with the size of w_{ii}

n: A vector of destination choice probability for inner-outer centroids with the size of w_{io}

q: A vector of destination choice probability for outer-inner centroids with the size of w_{oi}

r: A vector of destination choice probability for outer-outer centroids with the size of w_{oo}

P_{ii}: A link-OD use matrix for inner-inner OD pairs with the size of a x w_{ii}

P_{io}: A link-OD use matrix for inner-outer OD pairs with the size of a x w_{io}

P_{oi}: A link-OD use matrix for outer-inner OD pairs with the size of a x w_{oi}

P_{oo}: A link-OD use matrix for outer-outer OD pairs with the size of a x w_{oo}

$\mathbf{v}^*$: A vector of the observed traffic volume with the size of $\mathbf{a}$

Formulation

In this study, the traditional least square model proposed by Iida (2008) is adopted. This model minimizes the difference between observed and calculated link traffic volume considering the non-negativity constraints of traffic volume. The model estimates $\mathbf{g}$, generated traffic volume, and the inputs are the generated traffic volumes from the outside ($\mathbf{s}$) and link use ratio ($\mathbf{P}$). One of the practical advantages of this model is to explicitly consider the inflow from the outside. In the practical analysis, traffic flowing into or out of the study area is often set to be constant in order to keep the consistency among the surrounding areas, and the proposed model can handle this issue. The overall formulation of the OD estimation model can be written as follows;

[Objective function]

$$\min \Phi(\mathbf{g}) = \left(\left(\tau\mathbf{P}_{ii}^{T}\mathrm{diag}(\mathbf{m})\mathbf{D}_{ii} + (1-\tau)\mathbf{P}_{io}^{T}\mathrm{diag}(\mathbf{n})\mathbf{D}_{io}\right)\mathbf{g} + \left(\lambda\mathbf{P}_{oi}^{T}\mathrm{diag}(\mathbf{q})\mathbf{D}_{oi} + (1-\lambda)\mathbf{P}_{oo}^{T}\mathrm{diag}(\mathbf{r})\mathbf{D}_{oo}\right)\mathbf{s} - \mathbf{v}^{*}\right)^{T}\left(\left(\tau\mathbf{P}_{ii}^{T}\mathrm{diag}(\mathbf{m})\mathbf{D}_{ii} + (1-\tau)\mathbf{P}_{io}^{T}\mathrm{diag}(\mathbf{n})\mathbf{D}_{io}\right)\mathbf{g} + \left(\lambda\mathbf{P}_{oi}^{T}\mathrm{diag}(\mathbf{q})\mathbf{D}_{oi} + (1-\lambda)\mathbf{P}_{oo}^{T}\mathrm{diag}(\mathbf{r})\mathbf{D}_{oo}\right)\mathbf{s} - \mathbf{v}^{*}\right) \quad (1)$$

[Constraint]

$$\mathbf{g} \geq \mathbf{0} \quad (2)$$

Note that the operator diag(.) transforms the vector to the following matrix;

$$\mathrm{diag}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & x_n \end{pmatrix} \quad (3)$$

Let $\mathbf{c}_i$ and $\mathbf{c}_o$ be a vector of the traffic destined to inner and outer destinations respectively.

This can be expressed as follows;

$$c_i = Q_{ii}^{T}\mathrm{diag}(m)D_{ii}g + Q_{oi}^{T}\mathrm{diag}(q)D_{oi}s, \quad (4)$$

$$c_o = Q_{io}^{T}\mathrm{diag}(n)D_{io}g + Q_{oo}^{T}\mathrm{diag}(r)D_{oo}s. \quad (5)$$

Also define the summation of c_s.

$$c_i^a = \mathrm{sum}(c_i), c_o^a = \mathrm{sum}(c_o) \quad (6)$$

Using these, $\mathbf{m}$ and $\mathbf{q}$ can be expressed as follows.

$$m = \frac{Q_{ii}c_i}{c_i^a},\ q = \frac{Q_{oi}c_o}{c_o^a} \quad (7)$$

In short, the destination choice probabilities, $\mathbf{m}$ and $\mathbf{q}$, are the function of the traffic volume destined to the node, which is the function of $\mathbf{g}$. Therefore, $\mathbf{m}$ and $\mathbf{q}$ should satisfy the above recursive equations. This yields a very complex optimization problem. To simplify this, let us first assume that $\mathbf{m}$ and $\mathbf{q}$ are the constants. $\mathbf{m}$ and $\mathbf{q}$ will be updated in the outer loop, and the ordinary quadratic optimization problem will be solved in the inner loop.

2.3 Formulation of the inner loop algorithm

Define $\mathbf{A}$ and $\mathbf{b}$ as follows;

$$\mathrm{A:}\ \left(\tau \mathrm{P}_{ii}^{T}\mathrm{diag}(m)\mathrm{D}_{ii} + (1-\tau)\mathrm{P}_{io}^{T}\mathrm{diag}(n)\mathrm{D}_{io}\right) \quad (8)$$

$$\mathrm{b:}\ \left(\lambda \mathrm{P}_{oi}^{T}\mathrm{diag}(q)\mathrm{D}_{oi} + (1-\lambda)\mathrm{P}_{oo}^{T}\mathrm{diag}(r)\mathrm{D}_{oo}\right)s - v^{*} \quad (9)$$

Then the eq(1) can be written as follows.

$$\Phi(g) = (Ag+b)^{T}(Ag+b) \quad (10)$$

The problem has now become a simple quadratic optimization problem with non-negativity constraints.

$$\min\ \Phi(g) = (Ag+b)^{T}(Ag+b) \quad (11)$$

subject to $g \geq 0$

Outer loop algorithm

In the inner loop algorithm, the vectors $\mathbf{m}$ and $\mathbf{n}$ are treated as given values. However, as being easily imagined, these values cannot be predetermined and the accuracy of the estimation may rely heavily on these values. The outer loop is installed so as to update these values. Without any prior information, it may be reasonable to assume the equal destination choice probabilities. Alternatively if a small sample questionnaire survey can be conducted, we can start from the result of the survey. In any cases, the values of $\mathbf{m}$ and $\mathbf{n}$ are updated using the eqns (4)-(7).

VERIFICATION ON HYPOTHETICAL NETWORK

Network Overview

The shortage of the past OD estimation methods is that the integration of dummy traffic zone and actual road network is insufficient. In order to overcome this problem, a new idea for OD estimation method is proposed for OD estimation method. The numerical experiment is conducted in this section to verify our proposed approach. In this study the crossroad of two links can be seen as a centroid of a dummy zone in the hypothetical network. Thus, the hypothetical network can be consistent with actual road network. The relation between dummy traffic zone and road network is shown in Figure 1. The number of each link in hypothetical network is shown in Figure 2. In this hypothetical network, there are 6 inner nodes (No.1~6) and 4 outer nodes (No.11~14). Each link has two directions, which have identical travel time (see Figure 2 and 3). The basic setting of OD matrix is shown in Table 1.

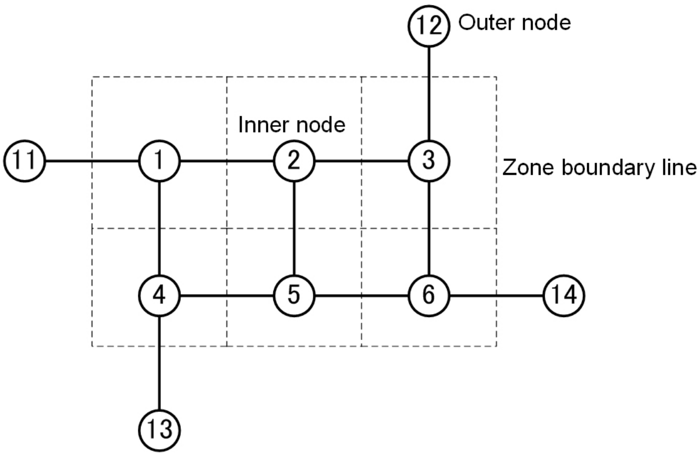

Figure 1 - Network and zone boundary

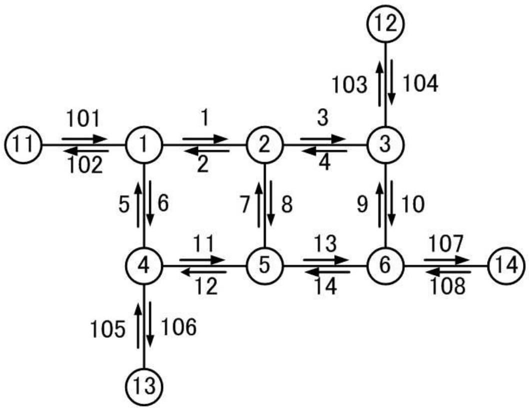

Figure 2 - Hypothetical network

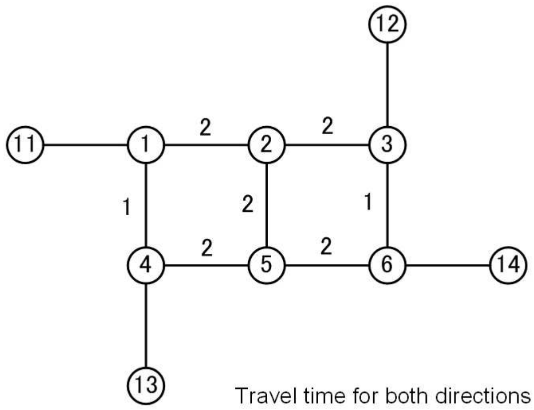

Figure 3 - Hypothetical link travel time

Table 1 - The hypothetical OD matrix

Original \ Destination		Inner Node						Sum of OD	Outer Node				Sum of OD	Total Sum of OD Production
		1	2	3	4	5	6		11	12	13	14		
Inner Node	1	0	176	80	64	208	128	656	50	60	40	50	200	856
	2	180	0	100	80	260	160	780	125	150	100	125	500	1280
	3	108	132	0	48	156	96	540	75	90	60	75	300	840
	4	90	110	50	0	130	80	460	75	90	60	75	300	760
	5	216	264	120	96	0	192	888	100	120	80	100	400	1288
Outer Node	6	162	198	90	72	234	0	756	75	90	60	75	300	1056
	11	72	88	40	32	104	64	400	0	120	80	100	300	700
	12	108	132	60	48	156	96	600	50	0	40	50	140	740
	13	90	110	50	40	130	80	500	125	150	0	125	400	900
	14	90	110	50	40	130	80	500	100	120	80	0	300	800
Total Sum of OD Attraction		1116	1320	640	520	1508	976	6080	775	990	600	775	3140	9220

Parameters representing ratio of traffic moved to inner centroids are τ and λ. In this case, τ is the rate of the traffic generated within the area destined within the area. If $\tau \to 0$, that means everything generated within the area goes out of the study area. λ is the rate of the traffic entered from outside destined within the area. If $\lambda \to 1$, that means everything comes into study area. τ and λ of each node can be calculated through Table 1. The calculation results of τ and λ are shown in Table 2. The average values of τ and λ are 0.6711 and 0.6369, respectively. In the following sections, the average values of τ and λ are treated as “the actual values” in the setting of numerical experiment.

Table 2 - The calculation results of τ and λ

—	**No.**	τ	$1-\tau$	λ	$1-\lambda$	**Average Value**
Inner Node	1	0.7664	0.2336	—	—	0.6711
	2	0.6094	0.3906	—	—	
	3	0.6429	0.3571	—	—	
	4	0.6053	0.3947	—	—	
	5	0.6894	0.3106	—	—	
	6	0.7159	0.2841	—	—	
Outer Node	11	—	—	0.5714	0.4286	0.6369
	12	—	—	0.8108	0.1892	
	13	—	—	0.5556	0.4444	
	14	—	—	0.625	0.3750	

We assume the logit-based decision making equation for creating the choice probability of each path between origin and destination. The formula of this index P_{oik} is :

$$P_{oik} = \exp(-\theta * T_{oik}) / \sum_k \exp(-\theta * T_{oik}) \tag{12}$$

where

P_{oik} : The choice probability of the *k*th path that locates between origin o and destination i.

T_{oik}: The travel time of the *k*th path that locates between origin o and destination i.

A parameter θ is assumed to be a controller of path choice probability in this numerical experiment. If θ approaches to infinity, the choice probability of the shortest path will arise dramatically. If θ approaches to 0, the choice probability of each path will be identical. The change of path choice probability under different settings of θ is listed in Table 3.

Table 3 - The change of path choice probability under different settings of θ

θ	**OD pair**	**Total travel time of each path**	**Path ID (The node in path)**	**Path choice probability**
1	2-4	3	1**(2→1→4)**	**0.7311**
		4	2**(2→5→4)**	**0.2689**
	1-6	5	1**(1→2→3→6)**	**0.4223**
		5	2**(1→4→5→6)**	**0.4223**
		6	3**(1→2→5→6)**	**0.1554**
0.5	2-4	3	1**(2→1→4)**	**0.6225**
		4	2**(2→5→4)**	**0.3775**
	1-6	5	1**(1→2→3→6)**	**0.3837**
		5	2**(1→4→5→6)**	**0.3837**
		6	3**(1→2→5→6)**	**0.2327**
0.1	2-4	3	1**(2→1→4)**	**0.525**
		4	2**(2→5→4)**	**0.475**
	1-6	5	1**(1→2→3→6)**	**0.3443**
		5	2**(1→4→5→6)**	**0.3443**
		6	3**(1→2→5→6)**	**0.3115**
0.01	2-4	3	1**(2→1→4)**	**0.5025**
		4	2**(2→5→4)**	**0.4975**
	1-6	5	1**(1→2→3→6)**	**0.3344**
		5	2**(1→4→5→6)**	**0.3344**
		6	3**(1→2→5→6)**	**0.3311**

Verification results

Evaluation index

The evaluation index we adopt to verify the accuracy of the proposed model is normalized root mean square error (NRMSE). The NRMSE can be expressed as follows:

$$\mathrm{NRMSE} = \sqrt{\sum_{\mathrm{i}} \left(\frac{\mathrm{Est_i} - \mathrm{Act_i}}{\mathrm{Act_i}}\right)^2 / \mathrm{N}} \tag{13}$$

$\mathrm{Est_i}$: The ith estimated value.

$\mathrm{Act_i}$: The ith corresponding actual value.

N : The number of cases.

The value of NRMSE can reflect the average level of accuracy because of the term, $\left(\frac{\mathrm{Est_i} - \mathrm{Act_i}}{\mathrm{Act_i}}\right)$. The value of NRMSE will be smaller when the estimated value is more close to the actual value. The estimated value is equal to the actual value when NRMSE=0. On the contrary, the value of NRMSE will be bigger when the estimated value is more farther from actual value.

Verification results of base case

In order to verify the performance of our proposed model, we conducted a test by using a small sized network (base case). The setting of the base case should be similar to the real situation. The predetermined conditions of the base case in this numerical experiment are as follows:

(1) τ and λ are 0.6711 and 0.6369, respectively.

(2) We assumed τ is consistent among nodes and so is λ.

(3) θ is assumed to be 0.1.

(4) The observed error of link traffic volume is assumed to be zero.

Under the predetermined conditions mentioned above, the verification results of the base case are shown in Figure 4 and Figure 5. Figure 4 is drawn to represent a relationship between the actual and the estimated link traffic volume. Figure 5 is drawn to represent a relationship between the actual and the estimated trip generation.

The results in both Figure 4 and Figure 5 seem to be reliable. The solid lines in Figure 4 and Figure 5 express the 45 degree line that mean the estimated values are completely equal to the actual values. In order to show a relationship between the estimated and actual values more clearly, the NRMSEs are calculated in terms of link traffic volume and trip generation, respectively. The NRMSE in terms of link traffic volume is **0.12**, and the NRMSE in terms of trip generation is **0.07**. From the viewpoint of NRMSE, we can conclude that the accuracy of estimated results is high.

In this section, the base case is utilized to verify the proposed approach. According to the results, we can conclude that the results of NRMSE in terms of link traffic volume and trip generation verify the performance of the proposed approach. Then, the higher the accuracy of input information is, the more precise the estimated link traffic volume is.

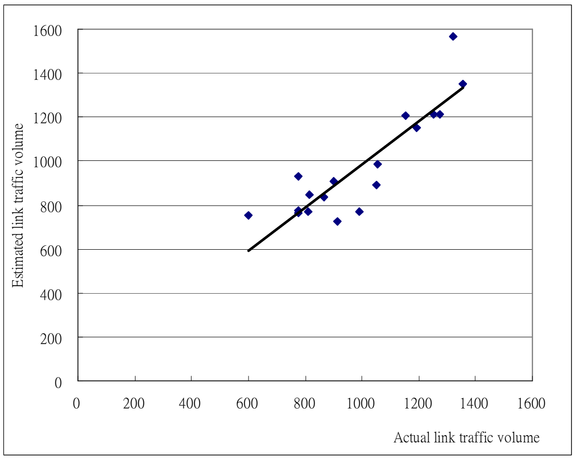

Figure 4 - The relationship between actual link traffic volume and estimated link traffic volume

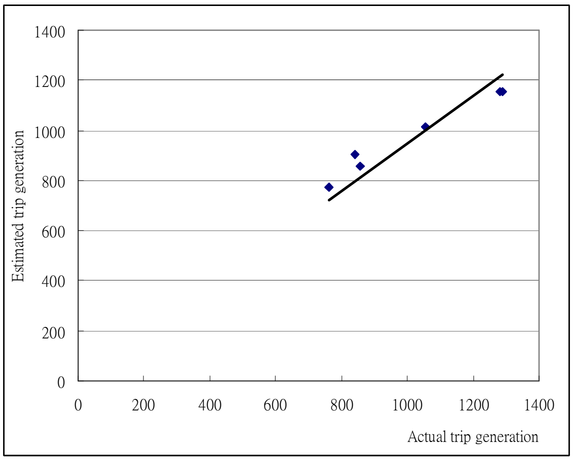

Figure 5 - The relationship between actual trip generation and estimated trip generation

In the inner loop algorithm, the vectors $\mathbf{m}$ and $\mathbf{n}$ are treated as given values. However, as being easily imagined, these values cannot be predetermined and the accuracy of the estimation relies heavily on these values. The outer loop is therefore installed so as to update these values. In order to confirm the efficiency of iteration, the sum of the square of the changes from the previous iteration is calculated. The result is as Figure 6. Judging from Figure 6, it is enough to iterate at most 10 times to have a stable result. In other words, this design of outer loop is useful for updating the value of $\mathbf{m}$ and $\mathbf{n}$ and only takes a little time on iteration.

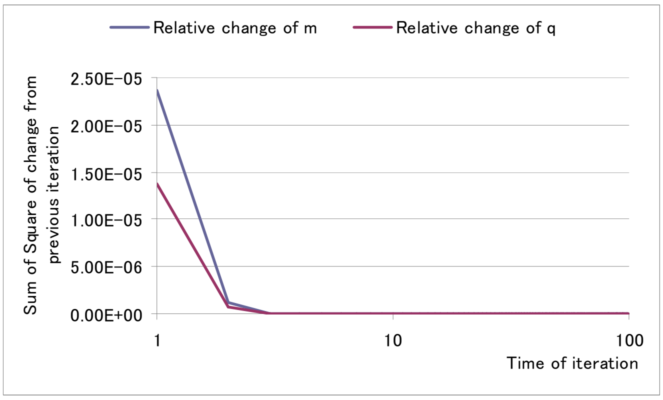

Figure 6 - Relative change of m and q under different time of iteration

Test of τ and λ

In general, we have some prior information on the parameters τ and λ, because the results of prior survey are available in most of cases. Accordingly it is assumed that the rough estimations of τ and λ are available for our computation, but we do not know their actual values exactly. In the numerical experiments for hypothetical network, we can use the "actual" values of OD demand, link traffic volumes, link usage proportion P_{oik} and the parameters τ and λ. It is necessary to assume the several cases of τ and λ, and to analyze the influence of these parameters on the accuracy of OD (or link traffic volume) estimation.

The influence of these parameters on the accuracy of OD estimation can be represented by using NRMSE between actual OD and estimated OD. The testing results of different cases of τ and λ are shown as Figure 7.

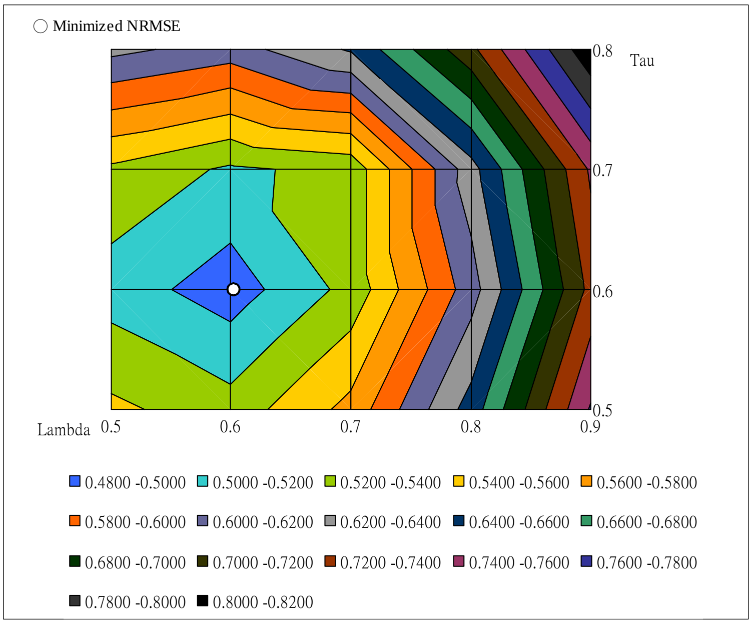

Figure 7 - NRMSE between actual and estimated link traffic volume

The influence of these parameters on the accuracy of link traffic volume estimation can be represented by using NRMSE between actual and estimated link traffic volume. The testing results of different cases of τ and λ are listed as Figure 8.

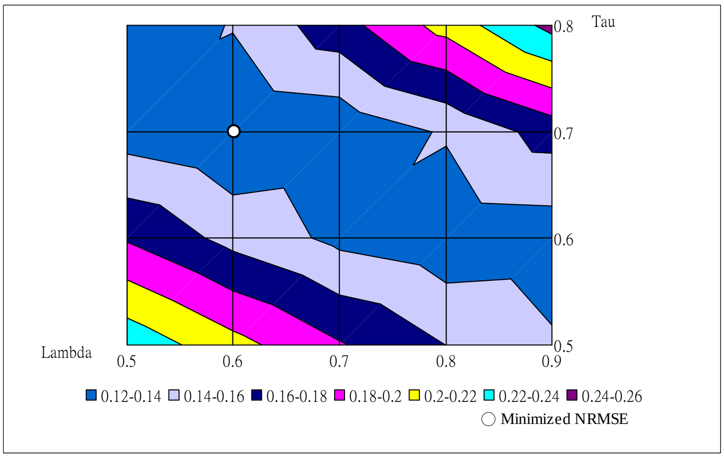

Figure 8 - NRMSE between actual and estimated OD

The results in Figure 7 and 8 both show how the level of estimation accuracy is influenced under different cases of τ and λ. We can obtain τ and λ close to the true value if we pick up the case with minimized NRMSE. In other words, if we know a little information about actual τ and λ, it may not be so hard to find optimal τ and λ through our approach. In this case, the optimal τ and λ are 0.7 and 0.6, respectively.

APPLICATION TO TAINAN NETWORK

Network overview

In this chapter, the proposed model for NCKU campus network is applied due to data availability. However, the accuracy of estimation is constrained, because it is very hard to obtain the actual OD data.

Figure 9 shows the actual road network and surrounding environment of NCKU campus. The part surrounding by dot lines is a study area for observation of link traffic volume. Figure 10 shows an abstract of NCKU campus network. There are 6 centriods (No. 1~6) where representing the inner traffic production/attraction. The outer nodes (No. 7~17) are used for representing outer traffic production/attraction. The remaining nodes represent actual crossroads or important entrances of campuses.

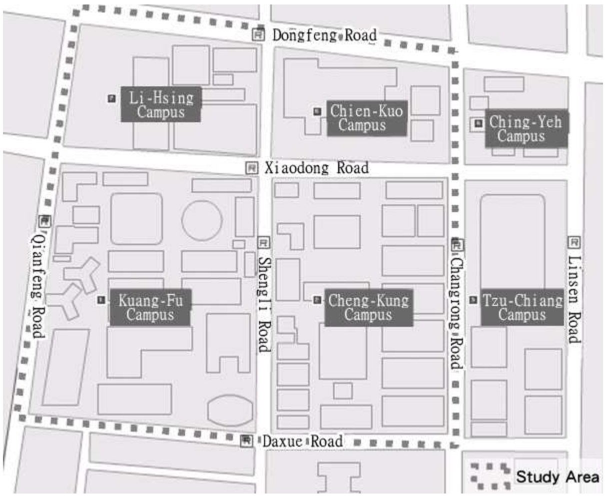

Figure 9 - NCKU campus and its surrounding roads

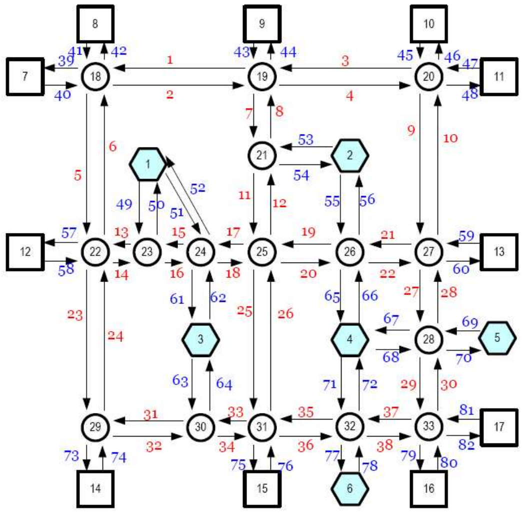

Figure 10 - NCKU campus network

4.2 Survey method

The proposed approach attempts to estimate OD traffic volume by using real link traffic volume only. Thus, the accurate observation of real link traffic volume becomes an important task before modelling. The conventional way of observing and recording link traffic volume is usually done by on-site counting. Surveyors are requested to stand on different corners of an intersection and record the corresponding link traffic volume by using counters. Four surveyors are needed in each intersection at the same time. Therefore, the total cost of survey is quite high. Moreover, the survey of every intersection cannot be done at the same time because of the limitation of manpower. Therefore, the problem of time lag usually exists in this kind of survey method. Generally speaking, the shortcomings of the conventional method are:

(1) Total cost of survey might be high,
(2) The quantity of surveyors might be large, and
(3) Total time of survey might be long.

Considering the shortcomings we mentioned above, digital cameras are used to record real link traffic volume at the same time. One digital camera could take care of one or two intersection at the same time. The digital cameras are set up on the top of buildings locating in the study area. The sites where digital camera locates are shown as star marks in Figure 11. 16 digital cameras were utilized and 12 surveyors were hired in this survey. The advantages of the new survey method are:

(1) Total cost of survey is lower than that of the conventional one,
(2) The quantity of surveyors is less than that of the conventional one,
(3) Total time of survey is shorter than that of the conventional one,
(4) The data collected by digital camera can be checked again and again for the purpose of accuracy, and
(5) The data can be utilized again even in different research.

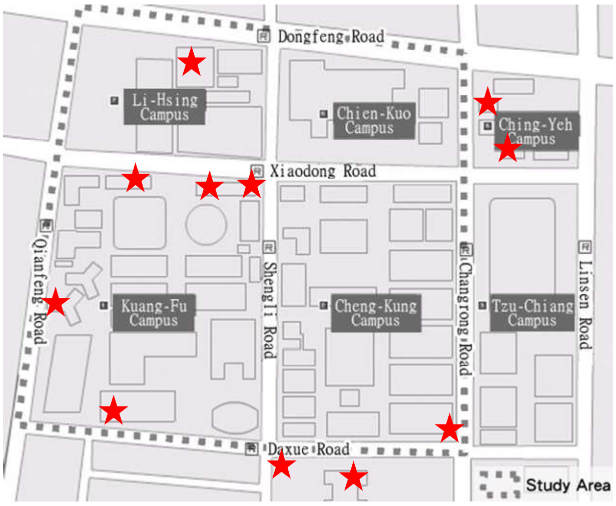

Figure 11 - The sites of camera setting

Observed variables

For the calculation of link usage rate, the length of each link was measured by using GIS. The shortest link length is 88.63 meters and the longest one is 522.17 meters. The links connected to inner and outer nodes (the link ID are 39 to 82, see Figure 10) are set to be 0. The result of measure is listed as Figure 12.

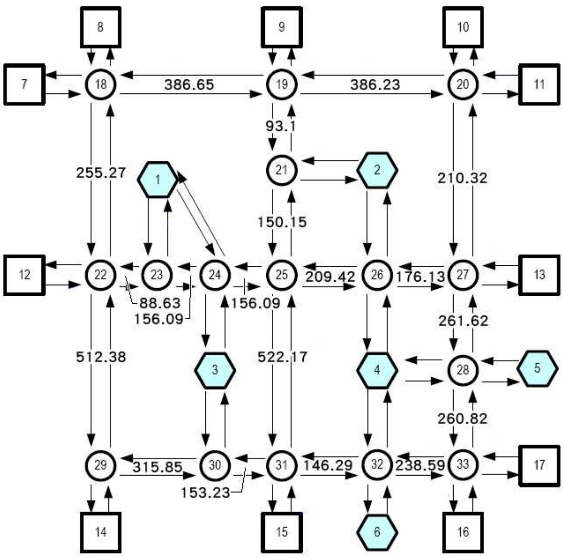

Figure 12 - Link distance

Note: The unit of measure is meter

For the purpose of OD estimation, the actual link traffic volume is needed to be observed. Therefore, cameras were set on the top of buildings. Link traffic volume was recorded during 10:00~10:15 and 15:00~15:15 through those cameras. Consequently, the observed link traffic volume in the morning and afternoon are shown as Figure 13 and 14 respectively. The periods of 10:00~10:15 and 15:00~15:15 are both not in peak period. The characteristics of real traffic flow during these two periods are relatively stable and constant. The traffic attracted into centroids is relatively small. The major part of traffic during these two periods is the through traffic.

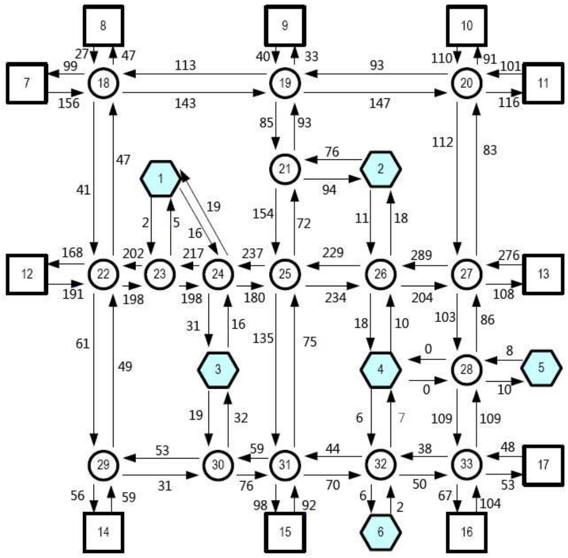

Figure 13 - Observed link traffic volumes in the morning

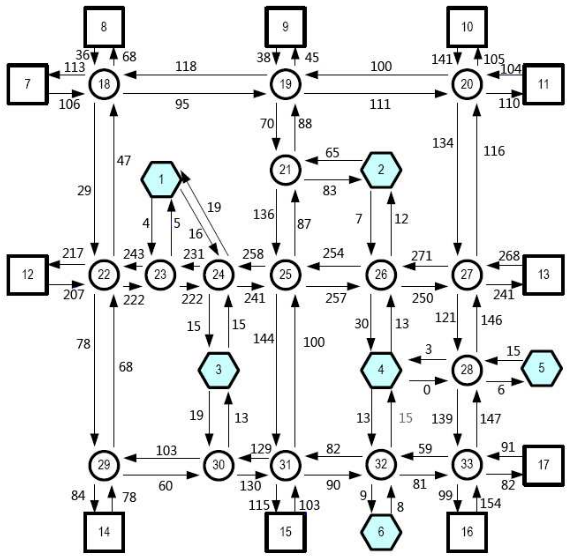

Figure 14 - Observed link traffic volumes in the afternoon

Estimation results

Due to the lack of true τ and λ in this survey, the trial-and-error process is adopted to investigate the optimal combination of τ and λ. The different pairs of τ and λ are set as inputs to obtain different OD traffic flow estimates. The final result of the morning model shows the optimal estimated result occurred when τ=0.04 and λ=0.35. In this case, τ is very small. That means the amount of traffic generated within the area destined within the area is quite few. The result is consistent with real situation because of the characteristics of the study area. The study area consists of four NCKU campuses. There are not so many trips moving from one campus to another. In addition, λ also gives us some information about the traffic entered from outside. Around one third of the traffic entered from the outside destined within the study area. The possible reason is that the NCKU hospital locates in study area. The traffic entered from outside mostly went to hospital rather than other campuses.

The final result of the afternoon model shows the optimal estimated result occurred when τ=0.15 and λ=0.38. The trips moving from one campus to another in the afternoon are more than in the morning. However, the value of λ is almost the same.

In order to clarify the accuracy of estimation, the relationship between the actual and the estimated link traffic volume in the morning is drawn in Figure 15 and that in the afternoon is drawn in Figure 16.

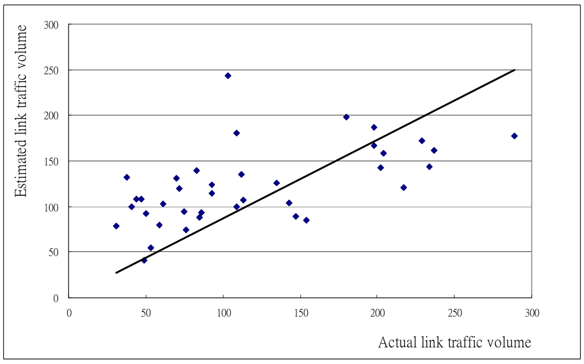

Figure 15 - The relationship between actual link traffic volume and estimated link traffic volume in the morning

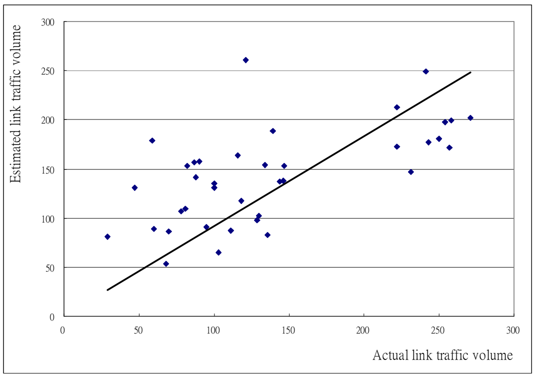

Figure 16 - The relationship between actual link traffic volume and estimated link traffic volume in the afternoon

As Figure 15 and 16 shown, the result is not so good, however, it verifies the estimation ability of the proposed approach. The reasons cause bad result of estimation might be:

(1) The assumption of consistent τ and λ for every centroid.
τ represents the rate of the traffic generated within the area destined within the area. λ represents the rate of the traffic entered from outside destined within the area. In an ideal case, τ and λ of each node can be calculated through trip-table. τ and λ do vary with centroids. However, the actual τ and λ cannot be obtained in this case due to the lack of actual trip-table information. Therefore, we assume they would not vary with centroids. That might cause some biases of estimation. In order to prevent the possible biases, the prior information on τ and λ is necessary. The car license plate tracing method is helpful for obtaining the information of τ and λ coordinating with sampling method.

(2) The actual path choice probability did not obtain in this survey.
Actually, the path choice probability is determined by distance factor only. Some factors that may affect route choice behaviour were not included in the algorithm. So the estimated path choice probability might deviate from actual one dramatically. That also

might cause some biases of estimation. Floating car survey or some detect methods will be necessary for obtaining real path choice probability.

Although our proposed approach is very simple and easy to operate, the prior information of τ, λ and path choice rate is hard to obtain. In the practice or future study, the efficient way of obtain the necessary information need to put more concern.

SUMMARY AND DISCUSIONS

This study establishes an OD flow estimation model from observed link traffic flow for solving problems traditional model faced and providing a practical tool for transportation planners. The advantages of the model developed in this research include: 1) the integration of dummy traffic zone and actual road network is sufficient ; 2) all the inner-inner, inner-outer, outer-inner and outer-outer OD traffic volume can be calculated at the same time; 3) the OD traffic volume is described as the function of the trip production only, and 4) the difference between the estimated and the measured link value of the traffic volumes measured from a link is minimized.

In the practice of urban transport planning especially for the OD estimation, the advantages of the model include: 1) this approach is relative simple and not complicated to operate. 2) only observed link traffic volume is needed for OD estimation. Coordinating with camera survey, time and cost of survey will be decreased. 3) the result of estimation can be easily understood because of the integration of dummy traffic zone and actual road network.

Although the estimation ability of the proposed model has been verified in real road network, it still remains some problems to improve. In future studies, the ITS technology can be considered to integrate into link traffic survey. It is helpful for improving the accuracy of observing and recording. In addition, Floating car survey or some detecting methods will be necessary for obtaining real path choice probability.

REFERENCES

Cascetta, E., Inaudi, D., and Marquis, G. (1993). Dynamic Estimators of Origin-Destination Matrices Using Traffic Counts, Transportation Science, Vol.27, No.4, 363-373.

Cremer, M. and Keller, H. (1987). A New Class of Dynamic Methods for the Identification of Origin-Destination Flows, Transportation. Res. –B, Vol. 21B, 117-132.

Davis, G A and Nihan, N L. (1991) Stochastic Process Approach to The Estimation of

Origin-Destination Parameters From Time Series of Traffic Counts, Transportation Research Record No. 1328, pp. 36-42.

Iida, Y. (2008) New Paradigm for Transport Planning, Gijutsu-Shoin.

Oneyama, H. and Kuwahara, M. (1997) Estimation of Time Department OD Matrices from Traffic Counts, Traffic Engineering, Vol. 32, No.2, 5-16 (in Japanese).

Sherali, H.D., Arora, N., and Hobeika, A.G. (1997) Parameter optimization methods for estimating dynamic origin-destination trip-tables, Transportation Research B, Vol.31, No.2, pp.141-157.

Sherali, H.D. and Park, T. (2001) Estimation of dynamic origin-destination trip tables for a general network, Transportation Research B, Vol.35, pp.217-235.

Sherali, H.D., Narayanan, A., and Sivanandan, R. (2003) Estimation of origin-destination trip-tables based on a partial set of traffic link volumes, Transportation Research B, Vol.37, pp.815-836.

Willumsen, L. G. (1984) Estimating Time-Dependent Trip Matrices from Traffic Counts, Ninth International Symposium on Transportation and Traffic Theory, VNU Science Press, 397-411.

Yang, H., Sasaki, T., Iida, Y. and Asakura, A. (1992) "Estimation of Origin-Destination Matrices from Link Traffic Counts on Congested Networks", Transportation Research, Vol.26B, No.6, pp.417-434.

Revision Report

Title : THE APPLICATION OF OD FLOWESTIMATION MODEL FROM OBSERVED LINK TRAFFIC FLOW: EMPIRICAL EVIDENCE IN TAINAN CITY

Hsia, Hao-Ching

Reviewer's opinion	Authors' Response	Revision
1) I am not an expert in the OD estimation area but find that very common papers by Sherali et al. are not reviewed or referenced in this paper. This is a huge concern for me. Are the authors aware of these papers or they decided to exclude them in the review. A strong reason is not provided for either case.	Thanks for the suggestions on literature review. Authors all aware of those efforts made by Sherali and his colleagues. The first reason for excluding them is the limitation of our paper size. The second one is the low relevance between our paper and Sherali and his colleagues' papers. Authors have modified some phrases to include the papers written by Sherali and his colleagues.	**Page 2. Line19.** For this reason, we are not looking at complex analytical models such as 'dynamic OD estimation (Sherali et al., 1997; Sherali et al., 2001; Willumsen, 1984; Oneyama and Kuwahara, 1997; Cascetta et al., 1993 and Cremer and Keller, 1987)', 'OD estimation with stochastic link traffic volumes (Davis and Nihan, 1991)', 'bi-level optimization formulation considering the congestion effect (Yang et al., 1994)', or 'sequential linear programming approach' (Sherali et al. 2003)
2) The paper presents results that do not seem to really enhance the results compared to traditional models. If this is the case reasons for lack of improvement need to discussed in detail	Authors thank for the opinion. We have made the reasons for lack of improvement more clearly in our paper. We also added some possible ways for solving problems in our proposed approach.	**Page 16. Line 7.** As Figure 15 and 16 shown, the result is not so good, however, it verifies the estimation ability of the proposed approach. The reasons cause bad result of estimation might be: (1) The assumption of consistent τ and λ for every centroid. τ represents the rate of the traffic generated within the area destined

		within the area. λ represents the rate of the traffic entered from outside destined within the area. In an ideal case, τ and λ of each node can be calculated through trip-table. τ and λ do vary with centroids. However, the actual τ and λ cannot be obtained in this case due to the lack of actual trip-table information. Therefore, we assume they would not vary with centroids. That might cause some biases of estimation. In order to prevent the possible biases, the prior information on τ and λ is necessary. The car license plate tracing method is helpful for obtaining the information of τ and λ coordinating with sampling method. (2) The actual path choice probability did not obtain in this survey. Actually, the path choice probability is determined by distance factor only. Some factors that may affect route choice behaviour were not included in the algorithm. So the estimated path choice probability might deviate from actual one dramatically. That also might cause some biases of estimation. Floating car survey or some detect methods will be necessary for obtaining real path choice probability. Although our proposed approach is very simple and easy to operate, the prior information of τ, λ and path choice rate is hard to obtain. In the practice or future study, the efficient way of obtain the necessary information need to put more concern.

3) Another very important concern of mine is the presence of many grammatical errors in the paper.	Authors thank for the opinion. We will check the paper again.	—
The applicability of the proposed system is demonstrated in a very simple scenario. However, the real world setting is much more complex. The paper will benefit by including a more realistic scenario. At the very least, the authors should discuss in more details potential benefits and pitfalls of using the proposed model to predict complex OD flows usually observed in the real life	1. Authors thank for the opinion. We demonstrate our proposed approach in a real network. However, we also need to make some assumptions for model calculation. We have described the assumptions we adopt in this paper clearly. Readers can understand the limitation of our proposed approach through our explanation. 2. The potential benefits of our proposed approach are listed in section 4.2.	As the first reviewer's question 2.