DYNAMIC ORIGIN–DESTINATION (O–D) MATRICES ESTIMATION FROM REAL-TIME TRAFFIC COUNT INFORMATION

Ofyar Z. TAMIN
Senior Researcher
Department of Civil Engineering
Institute of Technology Bandung
Jalan Ganesha 10, Bandung 40132
Fax: +62-22-2502350
E-mail: ofyar@trans.si.itb.ac.id

Hedi HIDAYAT
Senior Researcher
Department of Civil Engineering
Institute of Technology Bandung
Jalan Ganesha 10, Bandung 40132
Fax: +62-22-2502350
E-mail: hedi@trans.si.itb.ac.id

Ade SJAFRUDDIN
Senior Researcher
Department of Civil Engineering
Institute of Technology Bandung
Jalan Ganesha 10, Bandung 40132
Fax: +62-22-2502350
E-mail: ades@trans.si.itb.ac.id

Abstract: The conventional method to estimate O–D matrices requires very large surveys such as: home and roadside interviews; which are very expensive, labor intensive, and subject to large errors. Previous researches have been able to obtain the O–D matrices by using steady-state traffic counts information. ATCS (Area Traffic Control System) already installed in Bandung since 1997 provides us the real-time traffic count information for all signalized intersections. The technology for transferring data via Internet is also available and at a very low cost. The paper will explain the dynamic phenomena in estimating the O–D matrices in a real-time basis based on real-time traffic count information. This model can be used to study the day-to-day evolution of the O–D matrix relating to traffic flow fluctuation. Having known the real-time O–D matrix, several analysis and applications in a real-time basis can be conducted for solving urban transportation problems.

1. INTRODUCTION

Travel is an activity that has become part of our daily life and the demand for it always present problem especially in urban areas such as congestion, delay, air pollution, noise and environment. In order to alleviate these problems, it is necessary to understand the underlying travel pattern. The notion of Origin–Destination (O–D) Matrix has been widely used and accepted by transport planners as an important tool to represent the travel pattern. When an O–D is assigned onto the network, a flow pattern is produced. By examining this flow pattern, one can identify the problems that exist in the network and some kind of solution may be devised. An O–D matrix gives a very good indication of travel demand, and therefore, it plays a very important role in various types of transport studies, transport planning and management tasks.

Most techniques and methods for solving transportation problems (urban and regional) require O–D matrix information as a fundamental information to represent the transport demand. The conventional method to estimate O–D matrices requires very large surveys such as: home and roadside interviews; which are very expensive, lengthy, labor intensive, subject to large errors, and moreover, time disruptive to trip makers. As an illustration, for urban areas, the regional government of Jakarta can only afford to carry out this O–D survey three times during the last 23 years through very large and expensive transport projects such as: Jakarta Metropolitan Area Transportation Study (JMATS) in 1975, Arterial Road System Development Study (ARSDS) in 1987 and Transport Network Planning Regulation Study (TNPRS) in 1992.

Broad Outline of Nation's Direction (GBHN) 1993 has stated that all policies in transport development should be directed to perform an efficient, safe, comfort, reliable, and environmentally-based National Transportation System (Sistranas). The rapid changes in land use, population and employment, as well as vehicle ownership have resulted in the conventional methods are no longer suitable for developing countries. This is due to that the lengthy process (2–3 years) which will result in the information contained in the O–D matrices do not reflect anymore the real situation. Practically, it is frequently found that in solving the 1998 transportation problem, the 1992 O–D matrix is still being used due to the lack of information of the most recent O–D matrix information. Although the 1995 O–D survey had been carried out, however, the 1995 O–D matrix information is still not yet available.

All of these require an answer. Therefore, the new approach to tackle all of these problems is urgently required. The need for inexpensive methods, which require low-cost data, less time and less manpower generally called as 'unconventional method' is therefore obvious due to time and money constraint. This become even more valuable for problems which require 'quick-response' treatment such as urban transport problems due to high urbanization, rapid growth of population, improvement of income level, etc.

Traffic counts, the embodiment and the reflection of the O–D matrix; provide direct information about the sum of all O–D pairs which use those links. Some reasons why traffic counts are so attractive as a data base are: firstly, they are routinely collected by many authorities due to their multiple uses in many transport planning tasks. All of these make them easily-available. Secondly, they can be obtained relatively inexpensive in terms of time and manpower, easier in terms of organization and management and also without disrupting the trip makers. Therefore, a key element of the approach is a system to update the transport demand model using low-cost traffic count information.

2. METHODS FOR ESTIMATING AN O–D MATRIX

Methods for estimating an O–D matrix can be classified into 2 main groups as shown in figure 1. They are as follows: conventional and unconventional methods **(Tamin, 1988)**. Conventional methods rely heavily on extensive surveys, making them very expensive in terms of manpower and time, disruption to trip makers and most importantly the end products are sometimes short-lived and unreliable.

Another important factor is the complications that arise when following each stage of the modelling process. Furthermore, in many cases particularly in small towns and developing countries, planners are confronted with the task of undertaking studies under conditions of time and money constraints, which make the application of the conventional methods almost impossible. The introduction of inexpensive techniques for the estimation of O–D matrices will overcome the problem.

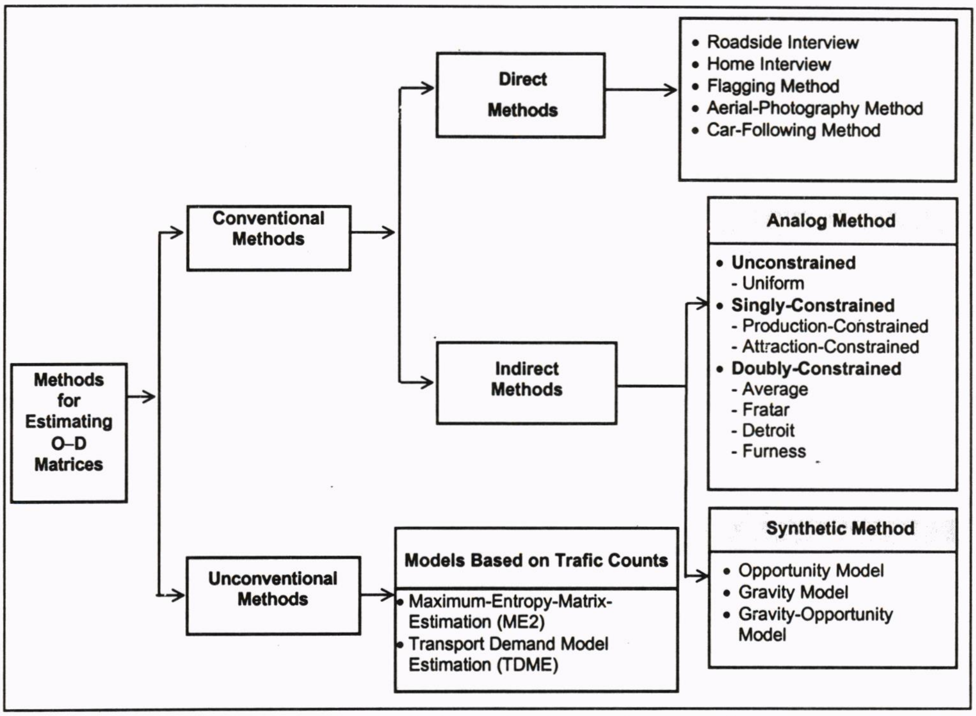

Figure 1. Methods for estimating an O–D matrix
Source: Tamin (1988)

As a result of dissatisfaction expressed by transport planners with conventional methods, other techniques for estimating O–D matrices which based on traffic counts have evolved over the years; these are generally called 'unconventional methods'. The aim of unconventional methods is to provide a simpler approach to solve the same problem and at a lower cost. Ideally, this simpler approach would treat the four-stage sequential model as a single process. To achieve this economic goal, the data requirements for this new approach should be limited to simple zonal planning data and traffic counts on some links or other low-cost data.

The first model based on traffic counts to be reported was probably developed by **Low (1972)**. The objective of his model was to 'effectively combines into one single process what is usually handled in a series of three or four sub-models, each with its own set of errors'. One of the advantages is that all the modelling errors appear in the final output in terms of traffic volumes and can be described statistically. The user thus has a better idea of how good his model is – something he does not know with the usual approach.

3. TRANSPORT DEMAND MODEL ESTIMATION BASED ON TRAFFIC COUNTS

3.1 General

Nguyen (1982) and **Tamin (1988,1999)** provides a very good and comprehensive overview on the state of art in this research domain related to the O–D matrix estimation based on traffic counts. They state the general problem in the following way. Let **P** denotes the sets of origins, **Q** denotes the set of destinations and **I=PxQ** denotes the set of origin–destination (O–D) pairs. Most of the existing models to estimate an O–D matrix $[T_{id}]$ from traffic counts may be written in the form:

$$\text{Minimum or maximum} \qquad S = f\left(\hat{V}_l, V_l\right) \tag{1}$$

$$\text{subject to} \qquad \sum_i \sum_d T_{id} \cdot p_{id}^l = \hat{V}_l \quad \text{for} \quad l \in \mathbf{L} \tag{2}$$

$$T_{id} \geq 0 \tag{3}$$

where: T_{id} = number of trips travelling from origin $\underline{i}$ to destination $\underline{d}$;
p_{id}^l = proportion of trips travelling from each origin $\underline{i}$ to each destination $\underline{d}$ that use link $\underline{l}$;
$\hat{V}_l, V_l$ = observed and estimated volume on link $\underline{l}$.

It can be seen that the value of p_{id}^l is defined by the route chosen by each user within the study area which can be estimated by applying suitable route choice technique. There are now available several route choice techniques ranging from the simplest one (all-or-nothing) to the most sophisticated one (equilibrium). Thus, theoretically, by knowing the information on $\hat{V}_l$ and p_{id}^l, the value of T_{id} can be estimated through mechanism of optimization equations (1)–(3).

3.2 Transport Demand Estimation Approach

The central idea is to develop estimation methods that can be used not only for estimating currently prevailing O–D matrix in a real-time basis and hence the O–D flows but also for forecasting O–D matrices and O–D flows which will prevail in the future. One possible way to develop methods for estimating O–D matrices from traffic counts is by modelling the trip making behaviour. The transport demand model estimation approach assumes that the travel pattern behaviour is well represented by a certain general transport model, e.g. a gravity model.

The main idea is to apply a system of transport models to represent the travel pattern. It should be noted here that the transport demand models are described as functions of some planning variables like population or employment and one or more parameters. Whatever the specification and the hypotheses underlying the models adopted, the main task is to estimate their parameters on the basis of traffic counts. Once, the parameters of the postulated transport demand models have been calibrated, they may be used not only for the estimation of the current O–D matrix, but also for predictive purposes. The latter requires the use of future values for the planning variables.

Consider a study area which is divided into $\underline{\mathbf{N}}$ zones, each of which is represented by a centroid. All of these zones are inter-connected by a road network which consists of series of links and nodes. Furthermore, the O–D matrix for this study area consists of $\underline{\mathbf{N}^2}$ trip cells. $(\underline{\mathbf{N}^2 - \mathbf{N}})$ trip cells if intrazonal trips can be disregarded. The most important stage is to identify the paths followed by the trips from each origin to each destination.

The variable p_{id}^{lk} is used to define the proportion of trips by mode $\underline{k}$ travelling from zone $\underline{i}$ to zone $\underline{d}$ through link $\underline{l}$. Thus, the flow on each link is a result of:

- trip interchanges from zone $\underline{i}$ to zone $\underline{d}$ or combination of several types of movement travelling between zones within a study area ($=T_{id}$); and
- the proportion of trips by mode $\underline{k}$ travelling from zone $\underline{i}$ to zone $\underline{d}$ whose trips use link $\underline{l}$, which is defined by p_{id}^{lk} $(0 \le p_{id}^{lk} \le 1)$.

The total volume of flow ($\hat{V}_l^k$) in a particular link $\underline{l}$ is the summation of the contributions of all trips interchanges by mode $\underline{k}$ between zones within the study area to that link. Mathematically, it can be expressed as follows:

$$V_l^k = \sum_i \sum_d T_{id}^k \cdot p_{id}^{lk} \tag{4}$$

Given all the p_{id}^{lk} and all the observed traffic counts ($\hat{V}_l^k$), then there will be $\underline{\mathbf{N^2}}$ unknown T_{id}^k's to be estimated from a set of $\underline{\mathbf{L}}$ simultaneous linear equations (1) where $\mathbf{L}$ is the total number of traffic (passenger) counts. In principle, $\underline{\mathbf{N^2}}$ independent and consistent traffic counts are required in order to determine uniquely the O–D matrix $[T_{id}^k]$. $\underline{(\mathbf{N^2{-}N})}$ if intrazonal trips can be disregarded. In practice, the number of observed traffic counts is much less than the number of unknowns T_{id}^k's.

3.3 Fundamental Basis

Using unconventional methods, it is assumed that the trip-making behaviour can be represented by a certain type of transport demand model such as gravity model. The link flows can then be represented as a function of a model form and its relevant parameters. The parameters of the postulated model are then estimated so that the errors between the estimated and observed traffic counts are minimized.

Consider now that there are $\underline{\mathbf{K}}$ trip purposes or commodities travelling between zones within the study area. Assume also that the interzonal movement within the study area can be represented by a certain transport demand model such as gravity (GR) model. Hence, the total number of trips T_{id} with origin in $\underline{i}$ and destination $\underline{d}$ for all trip purposes or commodities can be expressed as:

$$T_{id} = \sum_k T_{id}^k \tag{5}$$

T_{id}^k is the number of trips for each trip purpose or commodity $\underline{k}$ travelling from zone $\underline{i}$ to zone $\underline{d}$ as expressed by equation (6) generally known as a doubly-constrained gravity model (DCGR).

$$T_{id}^k = O_i^k . D_d^k . A_i^k . B_d^k \cdot f_{id}^k \tag{6}$$

A_i^k and B_d^k = balancing factors expressed as:

$$A_i^k = \left[\sum_d \left(B_d^k . D_d^k \cdot f_{id}^k\right)\right]^{-1} \text{ and } B_d^k = \left[\sum_i \left(A_i^k . O_i^k \cdot f_{id}^k\right)\right]^{-1} \tag{7}$$

$$f_{id}^k = \text{the deterrence function} = \exp\left(-\beta . C_{id}^k\right) \tag{8}$$

By substituting equation (6) to (4), the fundamental equation for estimating the transport demand model based on traffic counts is:

$$V_l^k = \sum_i \sum_d \left(O_i^k . D_d^k . A_i^k . B_d^k . f_{id}^k . p_{id}^{lk} \right) \tag{9}$$

The fundamental equation (9) has been used by many literatures not only to estimate the O–D matrices but also to calibrate the transport demand models from traffic count information (see **Tamin, 1988; Tamin and Willumsen, 1988**). Theoretically, having known the values of $\hat{V}_l^k$ and p_{id}^{lk}, T_{id}^k can be estimated by following the optimization mechanism of equations (1)–(3).

Equation (9) is a system of <u>L</u> simultaneous equations with only (<u>1</u>) unknown parameter β need to be estimated. The problem now is how to estimate the unknown parameters β so that the model reproduces the estimated traffic flows as close as possible to the observed traffic counts.

3.4 Estimation Methods

Tamin (1999) explains several types of estimation methods which have been developed so far by many researchers are:

- Least-Squares estimation method (LLS or NLLS)
- Maximum-Likelihood estimation method (ML)
- Bayes-Inference estimation method (BI)
- Maximum-Entropy estimation method (ME)

3.4.1 Least-Squares Estimation Method (LS)

Tamin (1988,1999) have developed several Least-Squares (LS) estimation methods of which its mathematical problem can be represented as equation (10).

to minimize $$S = \sum_l \left[\left(\hat{V}_l^k - V_l^k \right)^2 \right] \tag{10}$$

$\hat{V}_l^k$ = observed traffic flows for mode <u>***k***</u> V_l^k = estimated traffic flows for mode <u>***k***</u>

The main idea behind this estimation method is that we try to calibrate the unknown parameters of the postulated model so that to minimize the deviations or differences between the traffic flows estimated by the calibrated model and the observed flows.

Having substituted equation (9) to (10), the following set of equation is required in order to find an unknown parameter β which minimizes equation (10):

$$\frac{\delta S}{\delta \beta} = \sum_l \left[\left(2 \sum_i \sum_d T_{id}^k . p_{id}^{lk} - \hat{V}_l^k \right) . \left(\sum_i \sum_d \left(\frac{\partial T_{id}^k}{\partial \beta} . p_{id}^{lk} \right) \right) \right] = 0 \tag{11}$$

Equation (11) is an equation which has only one (1) unknown parameter β need to be estimated. Then it is possible to determine uniquely all the parameters, provided that L>1.

Newton–Raphson's method combined with the Gauss–Jordan Matrix Elimination technique can then be used to solve equation **(11)** (see **Batty, 1976; Wilson and Bennet, 1985**).

The LS estimation method can be classified into two: Linear-Least-Squares (LLS) and Non-Linear-Least-Squares (NLLS) estimation methods. **Tamin (1988)** has concluded that the NLLS estimation method requires longer processing time for the same amount of parameters. This may due to that the NLLS estimation method contains a more complicated algebra compared to the LLS so that it requires longer time to process. However, the NLLS estimation method allows us to use the more realistic transport demand model in representing the trip-making behaviour. Therefore, in general, the NLLS provides better results compared to the LLS.

3.4.2 Maximum-Likelihood Estimation Method (ML)

Tamin (1988,1999) have also developed an estimation method which tries to maximise the probability as expressed in equation **(12)**. The framework of the ML estimation method is that the choice of the hypothesis **H** maximising equation **(12)** subject to a particular constraint, will yield a distribution of V_l^k giving the best possible fit to the survey data ($\hat{V}_l^k$). The objective function for this framework is expressed as:

to maximize $$\mathbf{L} = \mathbf{c}.\prod_l p_l^{\hat{V}_l^k} \qquad (12)$$

subject to: $$\sum_l V_l^k - \hat{V}_T^k = 0 \qquad (13)$$

where: $\hat{V}_T^k$ = total observed traffic flows

$\mathbf{c}$ = constant

$p_l = \dfrac{V_l^k}{\hat{V}_T^k}$

By substituting equation **(9)** to **(12)**, finally, the objective function of ML estimation method can then be expressed as equation **(14)** with respect to unknown parameters β and θ.

$$\textbf{Max.} \quad \mathbf{L}_1 = \sum_l \left[\hat{V}_l^k .\log_e\left(\sum_i \sum_d T_{id}^k \cdot p_{id}^{lk} \right) - \theta . \sum_i \sum_d T_{id}^k \cdot p_{id}^{lk} \right] + \theta . \hat{V}_T^k - \hat{V}_T^k .\log_e \hat{V}_T^k + \log_e \mathbf{c} \qquad (14)$$

The purpose of an additional parameter θ, which appears in equation (14), is that to ensure the constraint equation **(13)** should always be satisfied. In order to determine uniquely parameter β of the GR model together with an additional parameter θ, which maximizes equation (14), the following two sets of equations are then required. They are as follows:

$$\frac{\delta \mathbf{L}_1}{\delta \beta} = \sum_l \left[\hat{V}_l^k . \frac{\sum_i \sum_d \frac{\delta T_{id}^k}{\delta \beta} \cdot p_{id}^{lk}}{\sum_i \sum_d T_{id}^k \cdot p_{id}^{lk}} \right] - \left(\theta . \sum_i \sum_d \frac{\delta T_{id}^k}{\delta \beta} \cdot p_{id}^{lk} \right) = 0 \qquad (15a)$$

$$\frac{\delta L_1}{\delta\theta} = -\theta.\left[\sum_i\sum_d T_{id}^k.p_{id}^{lk} - \hat{V}_T^k\right] = 0 \tag{15b}$$

Equation (**15ab**) is in effect a system of two (**2**) simultaneous equations which has two (**2**) unknown parameters β and θ need to be estimated. Again, the Newton–Raphson's method combined with the Gauss–Jordan Matrix Elimination technique can then be used to solve equation (**15ab**).

3.4.3 Bayes-Inference Estimation Method (BI)

The main idea behind the Bayes-Inference estimation method is by combining the prior beliefs and observations will produce posterior beliefs. If one has 100% confidence in one's prior belief then no random observations, however remarkable, will change one's opinions and the posterior will be identical to the prior beliefs. If, on the other hand, one has little confidence in the prior beliefs, the observations will then play the dominant role in determining the posterior beliefs. In other words, prior beliefs are modified by observations to produce posterior beliefs; the stronger the prior beliefs, the less influence the observations will have to produce the posterior beliefs. The objective function of the Bayes-Inference (BI) estimation method can be expressed as:

to maximize $$\mathbf{BI}\,(\tau_l^k V_l^k) = \sum_l \left(\hat{V}_l^k \log_e V_l^k\right) \tag{16}$$

By substituting equation (**9**) to (**16**), the objective function can then be rewritten as:

to maximize $$\mathbf{BI} = \sum_l \left[\hat{V}_l^k.\log_e\left(\sum_i\sum_d T_{id}^k.p_{id}^{lk}\right)\right] \tag{17}$$

In order to determine uniquely parameter β of the GR model, which maximizes equation (**17**), the following two sets of equations are then required. They are as follows:

$$\frac{\partial \mathbf{BI}}{\partial\beta} = \sum_l\left[\left(\frac{\hat{V}_l^k}{\sum_i\sum_d\left(T_{id}^k.p_{id}^{lk}\right)}\right).\left(\sum_i\sum_d\left(\frac{\partial T_{id}^k}{\partial\beta}.p_{id}^{lk}\right)\right) = 0\right] \tag{18}$$

Equation (**18**) is an equation which has one (**1**) unknown parameter β need to be estimated. Again, the Newton–Raphson's method combined with the Gauss–Jordan Matrix Elimination technique can then be used to solve equation (**18**).

3.4.4 Maximum-Entropy Estimation Method (ME)

Tamin (1998) has developed the maximum-entropy approach to calibrate the unknown parameters of gravity model. Now, this approach is used to develop procedure to calibrate the unknown parameters of the transport demand model based on traffic count information. The basic of the method is to accept that all micro states consistent with our information about macro states are equally likely to occur. **Wilson (1970)** explains that the number of micro states $W\{V_l^k\}$ associated with the meso state V_l^k is given by:

$$W[V_l^k] = \frac{V_T^k!}{\prod_l V_l^k!} \tag{19}$$

As it is assumed that all micro states are equally likely, the most probable meso state would be the one that can be generated in a greater number of ways. Therefore, what is needed is a technique to identify the values $[V_l^k]$ which maximize W in equation **(19)**. For convenience, we seek to maximize a monotonic function of W, namely $\log_e W$, as both problems have the same maximum. Therefore:

$$\log_e W = \log_e \frac{V_T^k!}{\prod_l V_l^k!} = \log_e V_T^k! - \sum_l \log_e V_l^k! \tag{20}$$

Using Stirling's approximation for $\log_e X! \approx X\log_e X - X$, equation **(20)** can then be simplified as:

$$\log_e W = \log_e V_T^k! - \sum_l \left(V_l^k \log_e V_l^k - V_l^k\right) \tag{21}$$

Using the term $\log_e V_T^k!$ is a constant; therefore it can be omitted from the optimization problem. The rest of the equation is often referred to as **the entropy function**.

$$\log_e W' = -\sum_l \left(V_l^k \log_e V_l^k - V_l^k\right) \tag{22}$$

By maximising equation **(22)**, subject to constraints corresponding to our knowledge about the macro states, enables us to generate models to estimate the most likely meso states (in this case the most likely V_l^k). The key to this model generation method is, therefore, the identification of suitable micro-, meso- and macro-state descriptions, together with the macro-level constraints that must be met by the solution to the optimisation problem. In some cases, there may be additional information in the form of prior or old values of the meso states, for example observed traffic counts ($\hat{V}_l^k$). The revised objective function becomes:

$$\log_e W'' = -\sum_l \left(V_l^k \log_e \left(\frac{V_l^k}{\hat{V}_l^k} \right) - V_l^k + \hat{V}_l^k \right) \tag{23}$$

Equation **(23)** is an interesting function in which each element in the summation takes the value zero if $V_l^k = \hat{V}_l^k$ and otherwise is a positive value which increases with the difference between V_l^k and $\hat{V}_l^k$. The greater the differences, the smaller the value of $\log_e W''$. Therefore, $\log_e W''$ is a good measure of the difference between V_l^k and $\hat{V}_l^k$. Mathematically, the objective function of the ME estimation method can be expressed as:

$$\textbf{to maximise} \quad E_1 = \log_e W'' = -\sum_l \left(V_l^k \log_e \left(\frac{V_l^k}{\hat{V}_l^k} \right) - V_l^k + \hat{V}_l^k \right) \tag{24}$$

In order to determine uniquely parameter β of the GR model which maximizes the equation (**24**), the following equation is then required. They are as follows:

$$\delta \mathrm{E}_1 = -\sum_l \left[\left(\sum_i \sum_d \frac{\partial T_{id}^k}{\partial \beta} \cdot p_{id}^{lk} \right) . \log_e \left(\frac{\sum_i \sum_d T_{id}^k \cdot p_{id}^{lk}}{\hat{V}_l^k} \right) \right] = 0 \qquad (25)$$

Equation (**25**) is an equation which has only one (**1**) unknown parameter β need to be estimated. Again, the Newton–Raphson's method combined with the Gauss–Jordan Matrix Elimination technique can then be used to solve equation (**25**).

3.4.5 Test Case With Steady State Traffic Count Data

The real data set of urban traffic movement in Bandung in terms of steady state traffic count information was used to validate the proposed estimation methods. Bandung is a capital of West Java Province and its population is around 6.4 millions in 1998 and expected to increase to 13.8 millions in 2020. The total area of Bandung is around 325,096 Ha and is divided into 66 kecamatans and 590 kelurahans.

The study area was divided into 146 zones of which 140 are internal zones and 6 are external. The road network of the study area consisted of 653 nodes and 1,811 road links. There are 95 observed 'steady-state' traffic counts ($\hat{V}_l$), traffic generation and attraction (O_i and D_d) for each zone, and observed O–D matrix for comparison purpose. The units used in equation (**9**) are as follows:

$\hat{V}_l$ = traffic counts in vehicles/hour
O_i, D_d = trip generation/attraction for each zone in vehicles/hour

The most important thing in 'transport demand model estimation from traffic counts' is to know how good the calibrated transport models are in reproducing the observed O–D matrix. There are two ways of doing this task:

a. the accuracy of the estimated O–D matrices compared to the observed one;
b. if the estimated O–D matrix is assigned onto the network then the corresponding traffic flows in each link should be as close as possible with the observed link flow obtained from ATCS control center.

In order to establish the strategy for validity and sensitivity tests, it is necessary to introduce at this stage the main issues affecting the accuracy of the estimated O–D matrix produced by the calibrated models. These are as follows:

- the choice of the transport demand model itself to be used in representing the trip making behaviour within the study area or, perhaps, a system of the real world;
- the estimation method used to calibrate the parameters of the transport model from traffic count information;
- number of traffic count information;
- the level of errors in traffic counts; and
- the level of resolution of the zoning system and the network definition.

The validity and sensitivity tests can then be established from these five main issues. Two transport demand models, namely gravity (GR) and gravity-opportunity (GO) models, and four estimation methods (NLLS, ML, BI, ME) have been used in the validity tests. The four estimation methods mentioned above have been discussed in detail in section 3.4.

The value of R^2 statistic as expressed in equation **(26)**–**(27)** is used to compare the observed and estimated O–D matrices to ascertain how close they are.

$$\mathbf{R}^2 = 1 - \frac{\sum_i \sum_d \left(\hat{T}_{id} - T_{id}\right)^2}{\sum_i \sum_d \left(\hat{T}_{id} - \mathrm{T}_1\right)^2} \tag{26}$$

$$\mathrm{T}_1 = \frac{1}{\mathrm{N}.(\mathrm{N}-1)} \cdot \sum_i \sum_d T_{id} \qquad \text{for } i \neq d \tag{27}$$

3.4.6 Important Findings

Several important findings can be concluded as given in table 1, which shows the performance ranking of model's estimation method according to specified criteria. The purpose of this table is to provide guidance to choose the best overall model's estimation method regarding its behaviour to several criteria such as: accuracy, computer time, sensitivity to errors in traffic counts, sensitivity to zoning level and network solution, and sensitivity to number of traffic counts.

The ranking scale ranging from 1 to 8 will be used to see the performance of estimation methods based on the above criteria. This approach is used to homogenise several types of quantitative scaling systems between each criteria into a 1–8 scaling system. Scale 1 shows the worst performance, while scale **8** shows the best performance.

Table 1. Performance ranking of model estimation methods for specified criteria

Model and estimation methods		Criteria				
		Accuracy	**Computer time**	**Sensitivity to errors in traffic counts**	**Sensitivity to zoning level and network resolution**	**Sensitivity to number of traffic counts**
GR	**NLLS**	6	8	8	7	4
	ML	2	6	7	8	3
	BI	1	6	6	5	1
	ME	3	5	5	6	2
GO	**NLLS**	8	4	NA	NA	8
	ML	5	2	NA	NA	7
	BI	4	3	NA	NA	5
	ME	7	2	NA	NA	6

Source: Analysis

It can be seen from table 1 that in terms of accuracy and sensitivity to number of traffic counts criteria, the GO model together with NLLS estimation performs the best. While, in terms of computer time, sensitivity to errors in traffic counts, sensitivity to zoning level and network resolution, the GR model with NLLS estimation performs the best. In general, it can be concluded that the NLLS estimation method shows the best ranking performance based on several types of criteria.

Table 2 shows the values of R^2 statistic of the observed O–D matrix compared with the estimated O–D matrices obtained from traffic counts.

Table 2. The value of R^2 for the comparison of the observed and estimated O–D matrices

Model	Estimation Methods				GR/GO[1]
	NLLS	ML	BI	ME	
GR	0.944	0.936	0.935	0.939	0.950
GO	0.946	0.943	0.942	0.945	0.956

Note: [1)] obtained using the observed O–D matrix information

Some final conclusions can then be drawn from table 2. They are as follows:

- in terms of O–D matrix level, it was found that the GO model always produced the best estimated matrices. However, these are only marginally better than those obtained by the GR model. Taking into account the results of using other criteria, it can be concluded that the best overall estimation methods are the combination of GR model with NLLS estimation method.

- with evidence so far, it was found that the estimated models and therefore O–D matrices are only slightly less accurate than those obtained directly from the full O–D surveys. This finding concludes that the transport demand model estimation approach is found encouraging in term of data collection and transport model estimation costs.

4. AREA TRAFFIC CONTROL SYSTEM (ATCS)

The Area Traffic Control System (ATCS) which have been installed in two large cities (Jakarta and Bandung) enable us to obtain the real-time traffic count information automatically for all signalized intersections. **DLGT (1996)** reports that the ATCS has been fully operated in Bandung since 1997. The technology for transferring data via Internet is also available and at a very low cost which enables us to obtain the traffic count information in a real-time basis.

Basically, the objective of ATCS deployment is to achieve the optimum traffic performance through minimization of intersection delay and creating continuous traffic flow called as green wave along the coordinated intersections. To achieve the above condition, the loop detectors record the traffic flow passing through the approaches. Then, the traffic data will be used for traffic signal arrangement interactively. The traffic data would be saved in the data base system at the ATCS control center through telecommunication network.

This traffic data is updated periodically in a real-time basis. The data base system can be accessed very easily at a very low cost through the Internet facility. This data would be as a main input data for dynamic O–D matrix estimation. As an illustration, Bandung has 117 intersections under ATCS and divided into two areas: the northern area consists of 59 intersections and the southern area consists of 58 intersections. The traffic data obtained from ATCS is traffic data in the approach of intersection. It is demanded to convert the data into link traffic data as required by the estimation process. This can be done through conversion factor.

5. THE CONCEPT IN DEVELOPING DYNAMIC AND REAL-TIME O–D MATRIX ESTIMATION MODEL BASED ON ATCS DATA

A systematic transport system modelling which try to make the best use of real-time traffic count information will be explained. The developed model can then obtain the real-time and dynamic O–D matrices relating to traffic flow fluctuation in a day in contrast to an ultimate steady-state traffic flow condition as being used by the conventional method. In general, the dynamic and real-time process for estimating the O–D matrices can be summarized by figure 2.

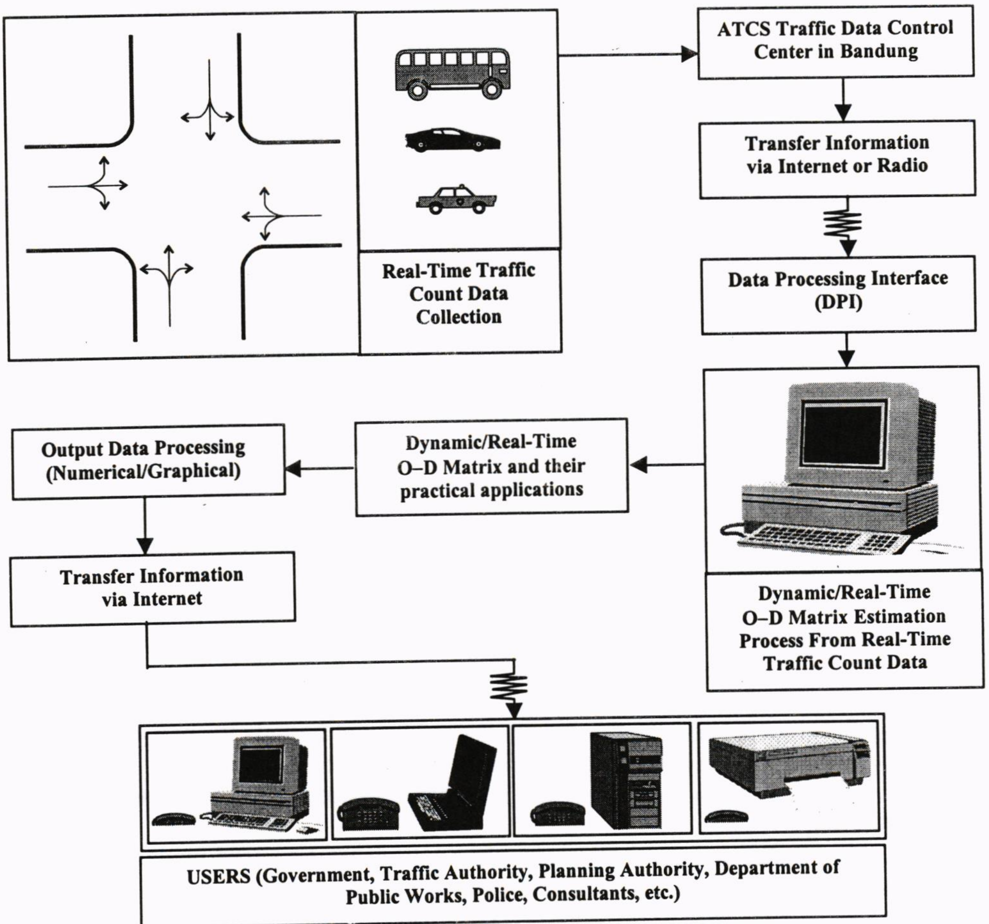

Figure 2. Process of dynamic/real-time O–D matrix estimation from real-time traffic count information

The real-time traffic data information is provided at the Traffic Control Center of ATCS project and can be directly and easily accessed using the Internet facility. Before the real-time traffic data is used in the O–D matrix estimation process; firstly, those data have to be processed in the Data Processing Interface (DPI). The process may include: error treatment due to transfer process, data formatting, data base preparation of zoning and network system, etc.

Having it processed; the traffic data will then be ready to be used for estimating the dynamic/real-time O–D matrices. The estimated dynamic/real-time O–D matrices and their practical applications will be stored in a Website so that the users can directly and easily access the information through the Internet facility. The Website is designed specifically and informatively for the purposes of user needs.

6. THE POTENTIAL OF DYNAMIC AND REAL-TIME O–D MATRICES IN SOLVING URBAN TRANSPORTATION PROBLEM

As mentioned above that some techniques and methods have been developed in very recent years which enable us to obtain the O–D matrices by using only easily available and low-cost traffic count information. The accuracy of estimated O–D matrices can reach up to 90% compared to those obtained by conventional methods. Unfortunately, at that time, the models still used the steady-state traffic count information obtained from the traffic count survey.

The latest development in automatic data collection for traffic count enables us to obtain the real-time traffic count information. For example, ATCS (Area Traffic Control System) already installed in Bandung since 1997 provides us the real-time traffic count information for all signalized intersections. Furthermore, the technology for transferring data is also readily available and at a very low cost through the use of Internet facility.

The use of real-time traffic count information enables us to analyze the dynamic phenomena of O–D matrices in a real-time basis. The developed model will give high added value through high efficiencies in terms of time and cost especially to be used to solve the dynamic and real-time urban transportation problem. In other words, we can obtain the accurate and low-cost O–D matrix information regularly within a very short period such as in every 1–2 hour.

Several things that have to be studied more carefully in order to increase the accuracy of the estimated O–D matrices are as follows:

a. development of the Data Processing Interface (DPI) and to study the best procedure for collecting real-time traffic count data from ATCS Control Center;
b. the conversion factor to convert the intersection-based traffic data into link-based traffic data;
c. better knowledge and obtaining more advanced transport demand models which will represent more accurately some specific travel demand pattern;
d. the optimum time-slice of real-time O–D matrices;
e. the optimum location and number of traffic count data and its impact to the accuracy of the estimated O–D matrices;
f. explanation on some unanswered questions relating to the impact of level of detail of zoning system and network definition on the real-time O–D matrices' accuracy;
g. more advanced route choice techniques (capacity-restrained or equilibrium) to take into account the effect of congestion especially in urban areas in relation to dynamic/real-time O–D matrix estimation from traffic counts;
h. the impact of the intersection delay to route choice and its effect to the accuracy of the estimated O–D matrices;

i. the evolution of real-time O–D matrices due to traffic flow fluctuation;

The output of real-time O–D matrices together with their practical applications will be stored in a Website designed specifically for the purposes of user needs (numerical and graphical). All users (Planning Authorities, Traffic Authorities, Department of Public Works, Consultants, Police, and other related agencies) can directly and easily access these informations at a very low cost through Internet facility.

Having known the real-time and dynamic O–D matrix, several analyses can be conducted and several applications can be carried out; some of them are:

a. to predict real-time O–D matrices based on fluctuated traffic, hence to provide the evolution of link flows as sources in identifying an appropriate road management scheme;
b. to provide real-time information on the performance of the network, both numerical and graphical e.g. link flows, link speeds, VCR values for all links, route guidance, locations of bottleneck, and many other real-time practical information;
c. to assess merits of the new introduction of new transport policy before it is implemented;
d. to analyze the effect of ATCS implementation on road traffic circulation;
e. several important applications which will solve the real-time urban transport problems.

7. CONCLUSIONS

The paper explains the dynamic phenomena in estimating the O–D matrices based on real-time traffic count information in which a novel unifying approach to describe the estimation of O–D matrices from traffic count information has also given. The significance of the model is, both theoretical and practical values; by understanding thoroughly the use of real-time traffic count information in obtaining the dynamic and real-time O–D matrices is a breakthrough giving high added value for applications in developing countries due to its effectiveness and efficient uses in many transport planning, engineering, management and policy tasks.

Such a model will enable us to obtain the dynamic O–D matrix (in a real-time basis) in contrast with the traditional, lengthy and costly O–D matrix obtained by home or roadside interview. This model can be used to study the day-to-day evolution of the O–D matrix relating to traffic flow fluctuation in a day in contrast to an ultimate steady-state traffic flow condition as being used by the traditional model. Moreover, this approach can be extended for regional applications.

The result of previous research, which utilized the steady-state traffic count information, was found very useful in developing the modified model based on real-time traffic count information. By using the real-time traffic count information, the dynamic and real-time O–D matrix can also be estimated. The real-time O–D matrix together with its applications will be then stored and provided in Website designed specifically for the purpose of users (numerically and graphically) so that it can be directly accessed and used by the users via internet at a very low cost.

Several important factors, which will strongly affect the accuracy of the estimated O–D matrices, are:

- the transport demand model itself in representing the trip making behaviour within the study area;
- the estimation method used to calibrate the model from traffic counts;
- trip assignment techniques used in determining the routes taken through the network;
- location and number of traffic counts;
- errors in traffic count information;
- finally, the level of resolution of the zoning system and the network definition.

REFERENCES

Batty, M. (1976) **Urban Modelling: Algorithm, Calibrations and Predictions**. Cambridge University Press, Cambridge.

DLGT (1996) **Bandung Area Traffic Control**, Final System Design, AWA Traffic System.

Low, D.E. (1972) A new approach to transportation systems modelling. **Traffic Quarterly, 26(3)**, 391–404.

Tamin, O.Z. (1988) The estimation of transport demand models from traffic counts. **PhD Dissertation**, University of London.

Tamin, O.Z. and Willumsen, L.G. (1988) Freight demand model estimation from traffic counts. **Proceedings of the 16th PTRC Summer Annual Conference**, Bath (UK).

Tamin, O.Z. and Soegondo, T. (1989) Freight demand modelling from traffic counts: a case study in Bali. **Proceedings of the 5th World Conference on Transport Research**, Yokohama (Japan).

Tamin, O.Z. (1990) Public transport demand estimation by calibrating a trip distribution–modal choice from traffic counts. **Proceedings of the 5th CODATU Conference**, Sao Paulo (Brazil).

Tamin, O.Z. (1992) Application of transport demand models for inter-regional vehicle movements in East Java. **Proceedings of the 6th World Conference on Transport Research**, Lyon (France).

Tamin, O.Z. (1998) The development of inference-bayes and maximum-entropy estimators for estimating transport demand models, Traffic Engineering Laboratory's Report, Department of Civil Engineering, ITB.

Tamin, O.Z. (1999) **Transport Planning and Modelling**, 2nd Edition. ITB Press, Bandung.

Wilson, A.G. (1970) **Entropy in Urban and Region Modeling**. Pion Ltd, London.

Wilson, A.G. and Bennet, R.J. (1985) **Mathematical Methods in Human Geography and Planning**. John Wiley and Sons Ltd.