THE DEVELOPMENT OF THE REAL TIME INTEGRATED TRAFFIC INFORMATION SYSTEM (RITIS) FOR INDONESIA

Ofyar Z. TAMIN

Professor
Traffic Engineering Laboratory, Department of Civil Engineering, Institute of Technology Bandung
Bandung, Indonesia

(Received May 31, 2001)

The paper is written based on research on 'Dynamic Origin-Destination Matrices From Real Time Traffic Count Information' completed in March 2001. The latest development in automatic traffic count data collection enables us to obtain the traffic count information in real time or short-time-interval basis. For example, ATCS (Area Traffic Control System) already installed in several large cities in Indonesia, such as: DKI-Jakarta (1994), Bandung (1997), and Surabaya (1998) provided us the short-time-interval traffic count information for all signalised intersections. This traffic data is updated periodically in a short-time-interval basis (e.g., 5, 15, or 30 minute time interval). This information is provided at the Traffic Control Centre (TCC) of ATCS project and can be directly and easily accessed at a very low cost through the Internet or a telephone line. This data is the main input for the short-time-interval OD matrix estimation. Before this traffic data is used in the OD matrix estimation process; firstly, these data have to be processed in the Data Processing Interface (DPI). Having it processed; the traffic data will then be ready for estimating the short-time-interval OD matrices.

The output of short-time-interval OD matrices together with several practical applications will be the main input for the Real Time Integrated Traffic Information System (RITIS). This information will be stored in a Website designed specifically and informatively for the purposes of user needs (numerical and graphical). This short-time-interval traffic system information will become the public-domain information which can be directly and freely accessed through the Internet by the users (e.g., planning authorities, traffic authorities, Department of Public Works, consultants, police, drivers, radio stations, and TV stations, and other related agencies, etc.).

One of the most important information is the best routes from each origin zone to each destination zone which have already considered the effect of congestion. This information will be the main data for the development of the Route Guidance System (RGS) so that each driver can choose his best route through the road network. The best route information will be changed in a short-time-interval basis depending on the traffic condition. Moreover, this approach can also be extended to provide the short-time-interval environmental information. The system has been tested and validated in Bandung and it showed remarkably good results for Bandung condition.

Key Words: Route guidance, Intelligent transport system, Urban transport, OD matrix, Real time traffic information

1. INTRODUCTION

As reported by Abidin[1], some papers concluded that most developed countries have faced losing billions of dollars in transportation just due to drivers not having enough prior information of what they thought was the best routes. Besides that, traffic congestion is blamed as the main factor which results in the loss of work productivity in USA around 100 billion dollars per year. It is also said that in 1991 there has been 40,000 fatal injuries in traffic accidents and more than 5 million injured. Moreover, the traffic accident has also contributed to losses as much as 70 billion dollars per year. Some other qualitative disadvantages are: delay, inefficient traffic movements, high fuel consumption due to congestion, air and noise pollution.

Travel is an activity that has become part of our daily life and the demand for it always presents problems especially in urban areas such as congestion, delays, air pollution, noise and environment. In order to alleviate these problems, it is necessary to understand the underlying travel pattern. As mentioned by Tamin[2], the notion of Origin-Destination (OD) matrix has been widely used and accepted by transport planners as an important tool to represent the travel pattern. When an OD matrix 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 OD 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 OD matrix information as fundamental information to represent the transport pattern. The conventional method to estimate OD matrices usually requires very large surveys such as: home and roadside interviews; which are very expensive, lengthy, labour intensive, subject to large errors, and

moreover, time disruptive to trip makers. As an illustration in Tamin[3], to obtain the national OD matrix, the Department of Transportation, Republic of Indonesia could only carry out this survey three times within a 25-year period (1982, 1988, and 1996).

Tamin et al.[4] mentioned that the broad outline of the 1998 Nation's Direction (GBHN) of Indonesia stated that all policies in transport development should be directed to perform an efficient, safe, comfortable, 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 being no longer suitable for developing countries. This is due to the lengthy process (2-3 years) which will result in the information contained in the OD matrices that do not reflect the real situation.

Practically, it is frequently found that in solving the 2001 transportation problem, the 1996 OD matrix is still being used due to the lack of the most recent OD matrix information. Although the 2001 OD survey is still undergoing, the 2001 OD matrix information will be available perhaps in the year 2003. Furthermore, during the monetary crisis, it is almost impossible to carry out this survey for the next 5-10 years. Moreover, for urban areas, the regional government of Jakarta can only afford to carry out this OD survey three times in the last 23 years as mentioned in Tamin et al.[4], 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.

All of these require an answer. This becomes 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. 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 an 'unconventional method' is therefore obvious due to time and money constraints.

Traffic counts, the embodiment and the reflection of the OD matrix; provide direct information about the sum of all OD pairs which use those links. Some reasons why traffic counts are so attractive as a database are: first, they are routinely collected by many authorities due to their multiple uses in many transport planning tasks. All of these make them easily available. Second, they can be obtained relatively inexpensively 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.

Previous research[2] and several other researchers have been able to obtain the OD matrices by only using traffic count information. Unfortunately, at that time, they still used the steady-state traffic count information obtained from the traffic count survey. Nowadays, the technology for automatic traffic data collection is very advanced. The latest development in automatic data collection for traffic count enables us to obtain the short-time-interval traffic count information.

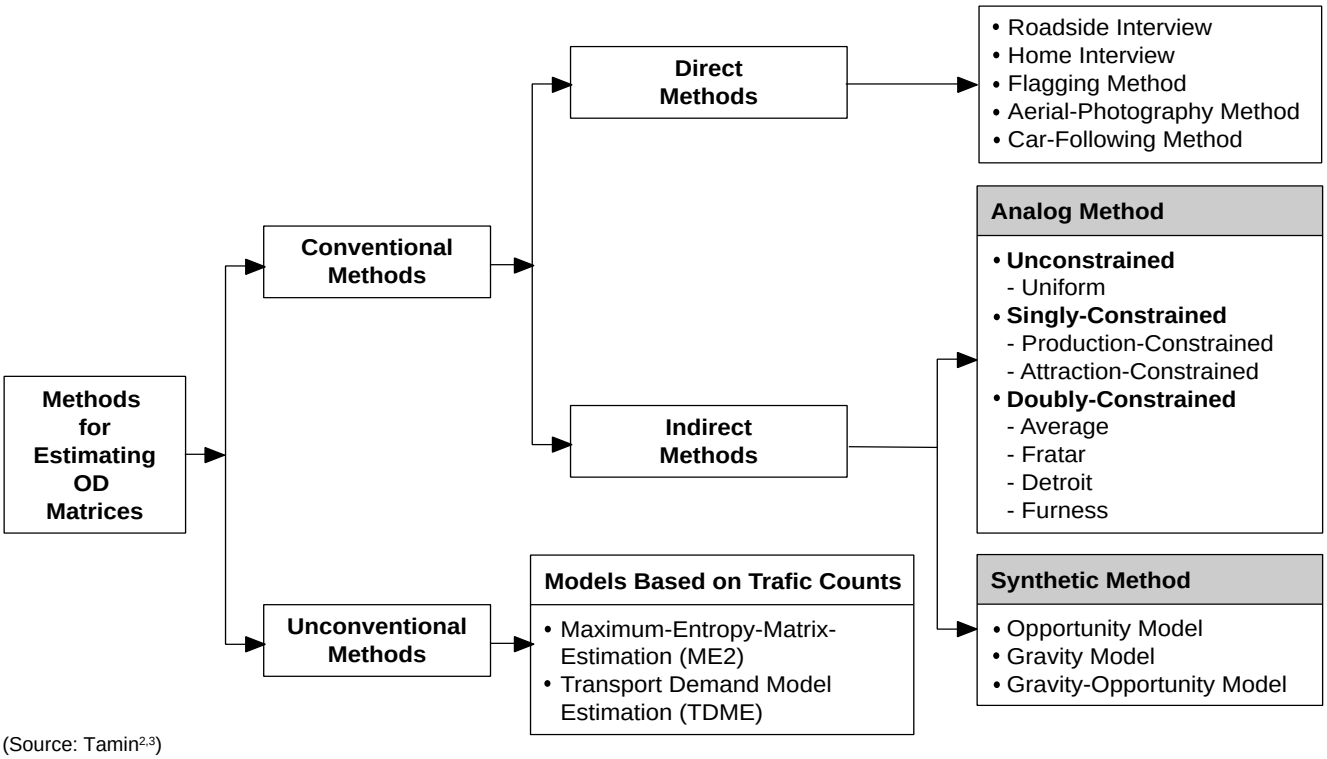
For example, as reported in AWA Plessey[5], ATCS (Area Traffic Control System) already installed in several large cities in Indonesia, such as: Jakarta (1994), Bandung (1997), and Surabaya (1998) provide us with the short-time-interval traffic count information for all signalised intersections. Furthermore, the technology for transferring data is also readily available and at a very low cost through the use of telephone lines. By using this information, the research is directed to develop a transport modeling system that enables us to produce the OD matrices in a short-time-interval basis.

2. METHODS FOR ESTIMATING OD MATRICES

Methods for estimating OD matrices can be classified into two main groups as shown in Figure 1.

They are as follows: conventional and unconventional methods[2,3]. 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 OD matrices will overcome the problem.

As a result of dissatisfaction expressed by transport



(Source: Tamin[2,3])

Fig. 1 Methods for estimating OD matrices

planners with conventional methods, other techniques for estimating OD matrices, which are 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.

There are several reasons why traffic counts are so attractive as a database to estimate the OD matrices:

a. Low-cost This type of data is relatively inexpensive to obtain since they require less manpower and automatic traffic counters can be used. They do not require preparation of questionnaires or statutory powers, and therefore they are easier in terms of organisation and management. They require simpler data analysis and output.

b. Availability Traffic counts are always available due to multiple uses in inter-urban or urban transport studies. They are widely used for different purposes like congestion analysis, accident studies, maintenance planning, intersection improvement, monitoring flow level and also used to determine expansion factors for OD surveys and to update OD matrices. Furthermore, many local authorities and planning bodies obtain these data regularly and hence the additional cost of using 'unconventional methods' is only marginal.

c. Non-disruptive Traffic counts can be obtained without generating any delay or disruption to vehicles. Furthermore, the automatic collection of traffic counts is well advanced and its accuracy is very satisfactory as well as there being several computer packages providing efficient processing.

Low[6] probably developed the first model based on traffic counts to be reported. The objective of his model was to 'effectively combine 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[7], Willumsen[8], and Tamin[3] provide a very good and comprehensive overview on the state of art in this research domain related to the OD matrix estimation based on traffic counts. They state the general problem in the following way. Let P denote the set of origins, Q denote the set of destinations and I = PxQ denote the set of origin-destination (OD) pairs. Most of the existing models to estimate an OD matrix $[T_{id}]$ from traffic counts may be written in the form:

$$\text{minimum or maximum} \quad S = f(\hat{V}_l, V_l) \quad \text{.....} (1)$$

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

$$T_{id} \geq 0 \quad \text{.....} (3)$$

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 $\underline{l}$, $0 \leq p_{id}^l \leq 1$;

$\hat{V}_l$, 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)[9]. To obtain the values of p_{id}^l, Tamin[2,3,10,13] has developed the procedure to estimate the values of p_{id}^l based on the predetermined trip assignment technique. Thus, theoretically, by knowing the information on $\hat{V}_l$ and p_{id}^l, the value of T_{id} can be estimated through the 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 OD matrix in a short-time-interval basis and hence the OD flows but also for forecasting OD matrices and OD flows which will prevail in the future. One possible way to develop methods for estimating OD 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[2,3]. 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 OD matrix, but also for predictive purposes. The latter requires the use of future values for the planning variables.

Consider a study area that is divided into $\underline{N}$ zones, each of which is represented by a centroid. All of these zones are inter-connected by a road network that consists of a series of links and nodes. Furthermore, the OD matrix for this study area consists of $\underline{N}^2$ trip cells. ($\underline{N}^2$–$\underline{N}$) 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} ($0 \leq p_{id}^{lk} \leq 1$).

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

$$V_l^k = \sum_i \sum_d T_{id}^k \cdot p_{id}^{lk} \quad \text{.....} (4)$$

Given all the p_{id}^{lk} and all the observed traffic counts ($\hat{V}_l^{\mathbf{k}}$), then there will be $\underline{N}^2$ unknown T_{id}^k's to be estimated from a set of $\underline{L}$ simultaneous linear equations (1) where L is the total number of traffic (passenger) counts. In principle, $\underline{N}^2$ independent and consistent traffic counts are required in order to determine uniquely the OD matrix $[T_{id}^k]$, ($\underline{N}^2$–$\underline{N}$) 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 a certain type of transport demand model such as the gravity model can represent the trip-making behavior. 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 $\underline{K}$ trip purposes or commodities travelling between zones within the study area. Assume also that a certain transport demand model such as gravity (GR) model can represent the interzonal movement within the study area. 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:

$$T_{id} = \sum_k T_{id}^k \quad \text{.....} \quad (5)$$

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).

$$T_{id}^k = O_i^k \cdot D_d^k \cdot A_i^k \cdot B_d^k \cdot f_{id}^k \quad \text{.....} \quad (6)$$

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

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

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

The readers who want to know more about the Gravity-Opportunity (GO) model are suggested to read Tamin[12] and Tamin and Soegondo[13]. By substituting equation (6) to (4), the fundamental equation for estimating the transport demand model based on traffic counts is:

$$V_l^k = \sum_i \sum_d (O_i^k \cdot D_d^k \cdot A_i^k \cdot B_d^k \cdot f_{id}^k \cdot p_{id}^{lk}) \quad \text{.....} \quad (9)$$

The fundamental equation (9) has been used by many papers not only to estimate the OD matrices but also to calibrate the transport demand models from traffic count information[2,3,10,11,12,13]. 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 $\underline{L}$ simultaneous equations with only one unknown parameter β that needs to be estimated. The problem now is how to estimate the unknown parameter β so that the model reproduces the estimated traffic flows as close as possible to the observed traffic counts.

3.4 Estimation methods

Tamin[3] explains several types of estimation methods that have been developed so far by many researchers as:

- 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[2, 3] have developed several Least-Squares (LS) estimation methods of which its mathematical problem can be represented as equation (10).

$$\text{to minimize} \quad S = \sum_l [(\hat{V}_l^k - V_l^k)^2] \quad \text{.....} \quad (10)$$

$\hat{V}_l^k$ = observed traffic flows for mode $\underline{k}$
V_l^k = estimated traffic flows for mode $\underline{k}$

The main idea behind this estimation method is that we try to calibrate the unknown parameters of the postulated model so as 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 equation is required in order to find an unknown parameter β which minimizes equation (10):

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

Equation (11) is an equation which has only one unknown parameter β that needs 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)[14,15].

The LS estimation method can be classified into two: Linear-Least-Squares (LLS) and Non-Linear-Least-Squares (NLLS) estimation methods. Tamin[2] has concluded that the NLLS estimation method requires a longer processing time for the same amount of parameters. This may be due to the NLLS estimation method containing a more complicated algebra compared to the LLS so that it requires a longer time to process. However, the NLLS estimation method allows us to use the more realistic transport demand model in representing the trip-making behavior. Therefore, in general, the NLLS provides better results compared to the LLS.

3.4.2 Maximum-Likelihood estimation method (ML)

Tamin[2,3] has also developed an estimation method that 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:

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

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

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

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.

$$\text{max. } L_1 = \sum_l \left[\hat{V}_l^k \cdot \log_e \left(\sum_i \sum_d T_{id}^k \cdot p_{id}^{lk} \right) - \theta \cdot \sum_i \sum_d T_{id}^k \cdot p_{id}^{lk} \right] + \theta \cdot \hat{V}_T^k - \hat{V}_T^k \cdot \log_e \hat{V}_T^k + \log_e c \quad \text{.....(14)}$$

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 set of equations are then required. They are as follows:

$$\frac{\delta L_1}{\delta \beta} = \sum_l \left[\hat{V}_l^k \cdot \frac{\sum_i \sum_d \frac{\delta T_{id}^k}{\delta \beta} \cdot p_{id}^{lk}}{\sum_i \sum_d T_{id}^k \cdot p_{id}^{lk}} \right] - \left(\theta \cdot \sum_i \sum_d \frac{\delta T_{id}^k}{\delta \beta} \cdot p_{id}^{lk} \right) = 0 \quad \text{.....(15a)}$$

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

Equation (15ab) is in effect a system of two simultaneous equations which has two unknown parameters β and θ that 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)

Tamin[3] has developed the Bayes-Inference (BI) estimation method in which the main idea is to combine the prior beliefs and observations to 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:

$$\text{to maximize } BI\,(T_l^k V_l^k) = \sum_l (\hat{V}_l^k \log_e V_l^k) \quad \text{.....(16)}$$

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

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

In order to determine the unique parameter β of the GR model, which maximizes equation (17), the following equation is then required:

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

Equation (18) is an equation which has one unknown parameter β that needs 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[5] has developed the maximum-entropy approach to calibrate the unknown parameters of the gravity model. Now, this approach is used to develop procedures to calibrate the unknown parameters of the transport demand model based on traffic count information. The basis of the method is to accept that all micro states consistent with our information about macro states are equally likely to occur. Wilson[16] explains that the number of micro states $W\{V_l^k\}$ associated with the meso state V_l^k is given by:

$$W\,[V_l^k] = \frac{V_T^k!}{\prod_l V_l^k!} \quad \text{.....(19)}$$

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:

$$\log_e W = \log_e \frac{V_T^k!}{\prod_l V_l^k!} = \log_e V_T^k! - \sum_l \log_e V_l^k! \quad ..(20)$$

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

$$\log_e W = \log_e V_T^k! - \sum_l (V_l^k \log_e V_l^k - V_l^k) \quad .....(21)$$

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.

$$\log_e W' = -\sum_l (V_l^k \log_e V_l^k - V_l^k) \quad ..... (22)$$

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:

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

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:

to maximise

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

In order to determine the unique parameter β of the GR model that maximizes the equation (24), the following equation is then required:

$$\delta E_1 = -\sum_l \left[\left(\sum_i \sum_d \frac{\partial T_{id}^k}{\partial \beta} \cdot p_{id}^{lk}\right) \cdot \log_e \left(\frac{\sum_i \sum_d T_{id}^k \cdot p_{id}^{lk}}{\hat{V}_l^k}\right)\right] = 0 \quad ...(25)$$

Equation (25) is an equation which has only one unknown parameter β that needs 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 the capital of West Java Province and its population is around 6.4 million in 1998 and expected to increase to 13.8 million by 2020. The total area of Bandung is around 325,096 hectares 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 OD 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 OD matrix. There are two ways of doing this task:

a. the accuracy of the estimated OD matrices compared to the observed one;
b. if the estimated OD 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 OD 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;
- location and 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 statistics as expressed in equation (26)–(27) is used to compare the observed and estimated OD matrices to ascertain how close they are[2,3].

$$R^2 = 1 - \frac{\sum_i \sum_d (\hat{T}_{id} - T_{id})^2}{\sum_i \sum_d (\hat{T}_{id} - T_1)^2} \quad \text{.....(26)}$$

$$T_1 = \frac{1}{N \cdot (N-1)} \cdot \sum_i \sum_d T_{id} \qquad \text{for } i \neq d \quad \text{.....(27)}$$

3.4.6 The best location of traffic counts

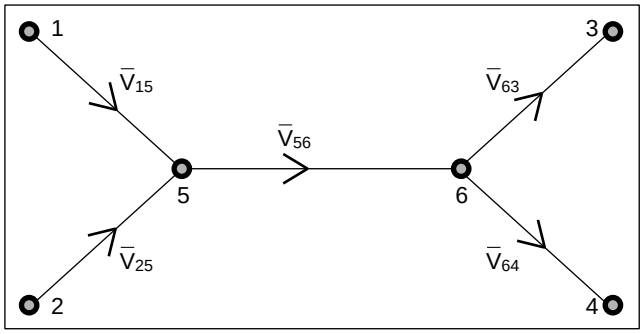
It is mentioned that the unconventional method uses traffic count information as the main input for estimating the OD matrices. Because of that, any process regarding the traffic counts should be clearly and deeply understood in order to obtain the best estimates of OD matrices; especially those which are related to data collection process, e.g., number of traffic counts and their best locations. The data collection process is very important since it is the first action in the whole process of OD matrix estimation. Some basic analyses used in finding the best location are as follows:

a. Proportion of trip interchanges on a particular link

The total volume of flow in a particular link $\underline{l}$ ($\hat{V}_l$) is the summation of the contributions of all trip interchanges between zones within the study area to that link. Mathematically, it can be expressed as equation (4). Tamin[3] stated that the most important stage for the estimation of OD matrix from traffic counts is to identify the paths followed by the trips from each origin $\underline{i}$ to each destination $\underline{d}$. In other words, the proportion of trip interchanges between zone $\underline{i}$ and zone $\underline{d}$ have to be uniquely identified for all those links involved. In this case, the variable p^l_{id} is used to define the proportion of trip interchanges from origin $\underline{i}$ to destination $\underline{d}$ travelling through link $\underline{l}$. Therefore, in finding the best location, the traffic counts having much information of the trip interchanges should be chosen. This information can be identified by analysing the total number and value of p^l_{id} in each link. This information will then be taken as the main criteria to determine the choice of the best location for traffic counts.

b. Inter-link relationship

- **Inter-dependence** Figure 2 shows that flows on link 5-6 are the summation of flows on link 1-5 and on link 2-5, then there is no additional information that can be extracted by counting flows on link 5-6 because of the flow continuity condition, $\overline{V}_{56} = \overline{V}_{15} + \overline{V}_{25}$.



(Source: Tamin[3])

Fig. 2 A simple network with link counts

In principle, we need counts for only three independent counts in order to find the flows of all links in Figure 2. Therefore, from an economic point of view, some efforts are needed in choosing the appropriate links to be counted[3].

- **Inconsistency** In practice, the problem of inconsistency in link counts may arise when the flow continuity conditions are not satisfied by the observed volumes. In the case of Figure 2, it may well happen that the observed flows are such that:

$$\overline{V}_{56} \neq \overline{V}_{15} + \overline{V}_{25} \quad \text{..... (28)}$$

or

$$\overline{V}_{15} + \overline{V}_{25} \neq \overline{V}_{63} + \overline{V}_{64} \quad \text{..... (29)}$$

This inconsistency in counts may arise due to human or counting errors and counting at different times or dates. As a result of all this, no solution for the OD matrix can be estimated that reproduces all these inconsistent traffic counts. One possible way to remove this problem is by choosing only independent links to be counted[3].

c. Optimum number of traffic counts

Equation (4) is the fundamental equation developed for estimating the OD matrices from traffic counts information. In this model, the parameters p^l_{id} are estimated using traffic assignment technique. Given all the p^l_{id} and all the observed traffic counts ($\hat{V}_l$), then there will be N^2 T_{id}'s to be estimated from a set of $\underline{L}$ simultaneous linear equations (1) where $\underline{L}$ is the total number of traffic counts. In principle, $\underline{N}^2$ independent and consistent traffic counts are required in order to determine uniquely the OD matrix $[T_{id}]$; $[\underline{N}^2-\underline{N}]$ if intrazonal trips can be disregarded. In practice, the number of observed traffic counts is much less than the number of unknowns T_{id}'s. Therefore, it is impossible to determine uniquely the solution[2].

d. The determination of the optimum number of traffic counts

As mentioned above, the determination of the optimum number of traffic counts will be conducted under one condition representing the sensitivity between the number of traffic counts and the link rank to the accuracy of the estimated OD matrices, namely: random condition. In this condition, several combinations of traffic counts will be created based on random selection. Each combination of traffic counts will then be used to estimate the OD matrices.

In this research, the initial OD matrix was created by calibrating the Gravity-Opportunity (GO) model from traffic counts by using all selected links (646 links). Tamin et al.[18] reports that the best of values of parameters (ε and μ) are ε = 0,4 and μ = 1,0 for the GO model. The other unknown parameters (α and β) of the GO model were then calibrated using 646 selected traffic counts by using Non-Linear-Least-Squares (NLLS) estimation method. The initial OD matrix to be used for comparison purposes will then be created using the GO model together with the values of its calibrated parameters. Figure 3 shows the relationship between the level of accuracy of the estimated matrices compared to the initial one and the number of selected traffic counts under random conditions.

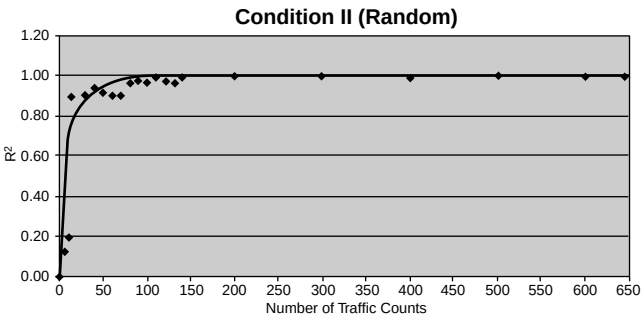


Fig. 3 Number of traffic counts and the value of R^2 (random condition)

It can be seen from Figure 3 that the use of 90 links has reproduced the relatively high accuracy of estimated OD matrices compared to the initial one (in terms of R^2). The use of 90 links has relatively the same accuracy with the use of 646 links. It can be concluded that the optimum number of traffic counts is 90 links (14% of 646 selected links or 3.6% of 2,485 available links).

3.4.7 Important findings

Table 1 shows the values of R^2 statistics of the observed OD matrix compared with the estimated OD matrices obtained from traffic counts.

Table 1 The value of R^2 for the comparison of the observed and estimated OD 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 OD matrix information

Some final conclusions can then be drawn from Table 1. They are as follows:

- In terms of OD 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 GO model with NLLS estimation method and the worst is GR model estimated by BI estimation method.
- With the evidence so far, it was found that the estimated models and therefore OD matrices are only slightly less accurate than those obtained directly from the full OD surveys. This finding concludes that the transport demand model estimation approach is found encouraging in terms of data collection and transport model estimation costs.

Several important findings can be concluded as given in Table 2, 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, sensi-

tivity to zoning level and network solution, and sensitivity to number of traffic counts.

Small differences of R^2 on Table 1 will then be regarded as reasonable differences to determine the superiority/inferiority among the estimation methods. The values of R^2 are transferred to the ranking scale ranging from 1 to 8 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.

It can be seen from Table 2 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.

4. AREA TRAFFIC CONTROL SYSTEM (ATCS)

4.1 Introduction

The Area Traffic Control System (ATCS) which has been installed in three large cities (Jakarta, Bandung, and Surabaya) enabled us to obtain the real-time or short-time-interval traffic count information automatically for all signalized intersections. DLGT[17] and AWA Plessey[5] report that the ATCS has been fully operational in Bandung since 1997. The technologies for transferring data via the Internet and telephone lines are also available and at very low cost that enables us to obtain the traffic count information in a short-time-interval basis. Basically, the objective of ATCS is to achieve the optimum traffic performance through minimization of intersection delay and creating continuous traffic flow called a 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 database system at the ATCS control center through telecommunication network.

This traffic data is updated periodically in a short-time-interval basis. The database system can be accessed very easily at a very low cost through the Internet and telephone line facilities. This data would be the main input data for short-time-interval OD matrix estimation. As an illustration, as reported in AWA Plessey[5], 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 required to convert the data into link traffic data as required by the estimation process. This can be done through the conversion factors.

4.2 The development of Data Processing Center (DPC)

As we know that the short-time-interval data traffic count will be used as the main data. This system has some sensors (detectors) in every direction point and in general terms it can be seen in Figure 4.

For the final process, the traffic count files are required to estimate the OD matrices. These files are taken

Table 2 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

Note : *) concluded from Table 1
Source : Analysis

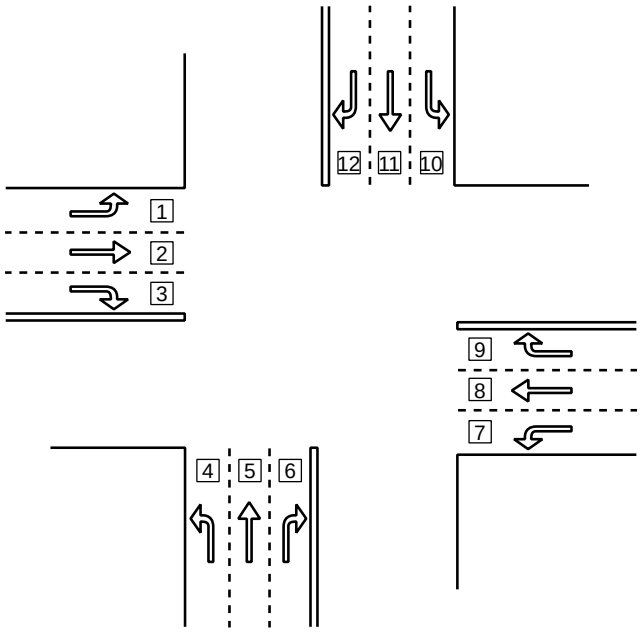


Fig. 4 The general terms of sensors in a signalised intersection

periodically from ATCS. All the traffic data of every sensor in each intersection are sent to the ATCS server and saved into a file which have to be ordered first. In ordering the file, some variables should be specified such as: location of intersection, and the saving period (date, hour, minute). Another problem is the data communication. Some steps have to be followed:

- **developing data communication from ATCS server to laboratory** The data communication required is something reliable and not too expensive. After some analysis, the telephone line medium is chosen which is quite cheap especially in operational costs.
- **developing communication to the ATCS server** The operating system in ATCS server is quite different with other well-known operating systems. It uses the OpenVMS of Digital/DEC.

For those two steps above, an attached computer facility is required. The main purpose of this system is to provide data in short-time-interval basis; therefore, a time-basis process is required. The automation process includes: process in ordering the data, taking data from the ATCS server, sending data using a telephone line, converting data from VS formatted into text formatted, processing data into MOTORS package program, and, saving it into a documentation file. In this step, the text-formatted files need to be converted into a file that can be used by MOTORS package program.

5. THE CONCEPT IN DEVELOPING THE REAL TIME INTEGRATED TRAFFIC INFORMATION SYSTEM (RITIS)

5.1 Basic configuration

In general, the basic configuration of the system is summarized in Figure 5. Collecting automatically the traffic count information obtained from the Traffic Control Center (TCC) of ATCS in a short-time-interval basis develops this information system. However, some important processes are required such as data transferring, data processing, and output processing, as follows:

- **Data transferring:** Traffic counts at signalised intersections were detected via wire detectors and automatically connected to the Traffic Control Center (TCC) of ATCS. This traffic data is updated periodically in a short-time-interval basis. The traffic data information is then provided at the TCC and can be directly and easily accessed at a very low cost via a telephone line.
- **Data processing:** This data is the main input for short-time-interval OD matrix estimation. Before the traffic data is used in the estimation process; firstly, those data have to be processed in the Data Processing Center (DPC). The process includes: error treatment due to transfer process, data formatting, database preparation of zoning and network system, etc. Moreover, the traffic data obtained from ATCS is traffic at 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. Having it processed, the traffic data will then be ready to be used for estimating the short-time-interval OD matrices.
- **Output processing:** The estimated short-time-interval OD matrices and their practical applications are stored in a Real Time Integrated Traffic Information System (RITIS) so that the users can directly and easily access the information through the Internet facility. The RITIS is designed specifically and informatively for the purposes of user needs.
- **Route information processing:** The driver can contact the Data Processing Center (DPC) to inform his location and his destination zone in which his location will be determined by GPS. The DPC will then process the best route by considering the latest traffic conditions and send the best route information back to the vehicle.

All of these processes will be designed in the forms of Route Guidance System (RGS) and the Real Time Integrated Transportation Information System (RITIS).

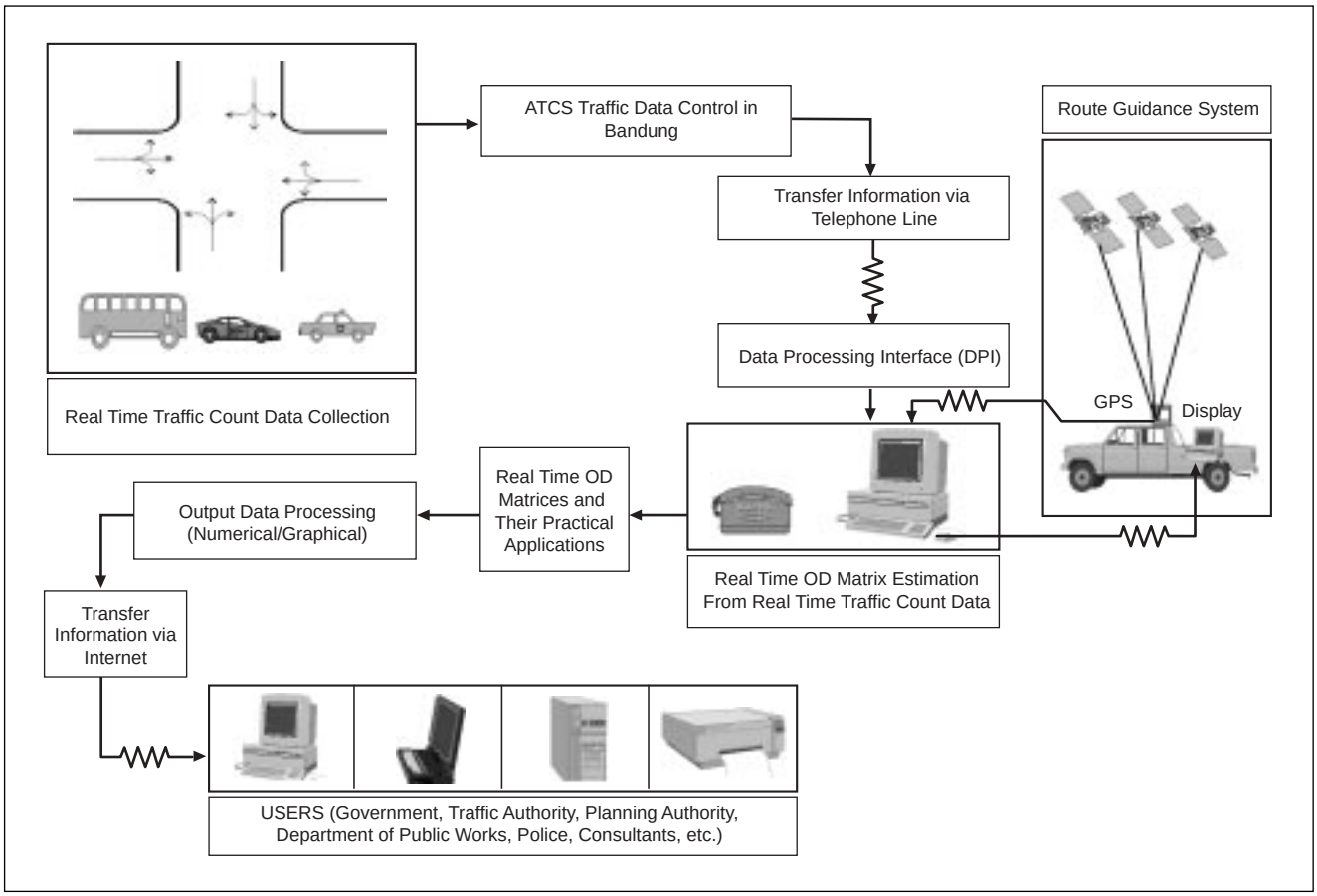


Fig. 5 The basic configuration of Real Time Integrated Traffic Information System (RITIS)

6. THE POTENTIAL OF RITIS IN SOLVING URBAN TRANSPORTATION PROBLEMS

As mentioned above, some techniques and methods have been developed in very recent years which enable us to obtain the OD matrices by using only easily available and low-cost traffic count information. 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 counts enables us to obtain the short-time-interval traffic count information. For example, ATCS (Area Traffic Control System) already installed in Bandung since 1997 provides us with real-time or short-time-interval 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 the Internet or a telephone line facility.

The use of short-time-interval traffic count information enables us to analyze the dynamic phenomena of OD matrices in a short-time-interval basis. The developed model will give high added value through high efficiencies in terms of time and cost especially to be used to solve dynamic urban transportation problems. In other words, we can obtain the accurate and low-cost OD matrix information regularly within a very short period such as every 30 minutes.

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

a. development of the Data Processing Interface (DPI) and to study the best procedure for collecting short-time-interval 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 trans-

port demand models that will represent more accurately some specific travel demand patterns;

d. the optimum time-slice of OD matrices;
e. the optimum location and number of traffic count data and its impact on the accuracy of the estimated OD matrices;
f. explanation on some unanswered questions relating to the impact of level of detail of zoning system and network definition on the OD 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 OD matrix estimation from traffic counts;
h. the impact of the intersection delay to route choice and its effect to the accuracy of the estimated OD matrices;
i. the evolution of short-time-interval OD matrices due to traffic flow fluctuation.

The output of short-time-interval OD matrices together with their practical applications will be stored in a Real Time Integrated Traffic Information System (RITIS) 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 this information at a very low cost through Internet facilities.

Several transport analyses can be conducted and several applications can be carried out by using the RITIS, some of them are:

- to predict short-time-interval (30 minute) OD matrices based on fluctuating traffic, hence to provide the evolution of link flows as sources in identifying an appropriate road management scheme;
- to provide short-time-interval information on the performance of the network, both numerical and graphic, e.g., link flows, link speeds, VCR values for all links, route guidance, locations of bottlenecks, and many other short-time-interval practical information;
- to assess merits of the introduction of new transport policy on the road network performance before it is implemented;
- to analyze the effect of ATCS implementation on road traffic circulation;
- several important applications which will solve dynamic urban transport problems.

This short-time-interval traffic information will become the public-domain information that can be directly and freely accessed via the Internet by users (e.g., road users, traffic police, traffic planners, traffic authority, radio stations and TV stations, etc.). Moreover, this approach can also be extended to provide the short-time-interval environmental information. This method has been tested and validated in Bandung and it shows remarkably good results for Bandung conditions. Several further applications can also be developed such as: the route guidance system for private vehicle and taxi, bus operating system, fleet management, etc.

7. CONCLUSIONS

The paper explains briefly the concept of developing the Real Time Integrated Traffic Information System (RITIS) and the Route Guidance System (RGS) by utilizing the short-time-interval traffic count information. A novel unifying approach to describe the estimation of OD matrices from traffic count information has been given. The significance of the model is, both theoretical and practical values; by understanding thoroughly the use of short-time-interval traffic count information in obtaining the short-time-interval OD matrices is a breakthrough giving high added value for applications in developing countries due to its effectiveness and efficient uses in transport planning, engineering, management and policy tasks.

The result of previous research, which utilised the steady-state traffic count information, was found very useful in developing the modified model based on short-time-interval traffic count information. By using this traffic count information, the short-time-interval OD matrix can also be estimated. The OD matrices together with their applications will be stored and provided in the RITIS designed specifically for the purposes of users (numerically and graphically) so that it can be directly accessed and used via the Internet at a very low cost.

One of the most important pieces of information is the best routes from each origin zone to each destination zone which has already considered the effect of congestion. This information will be the main data for the development of the Route Guidance System (RGS) so that each driver can choose his best route through the road network. The best route information will be changed in a short-time-interval basis depending on the traffic condition. Moreover, this approach can also be extended to provide short-time-interval environmental information.

REFERENCES

1. Abidin. The potential and prospect of GPS for transportation sector in Indonesia. Proceedings of 1st Symposium of Indonesian Inter-University Transport Studies Forum (IIUTSF), 3 December 1998, Bandung (in Indonesian). (1998).
2. Tamin, O.Z.The estimation of transport demand models from traffic counts. Ph.D. Dissertation, University of London. (1988).
3. Tamin, O.Z. Transport Planning and Modelling, 2nd Edition. ITB Press, Bandung. (2000).
4. Tamin, O.Z., Hidayat, H., Sjafruddin, A. The concept of developing the real time OD matrices based on ATCS data. Proceedings of 1st Symposium of Indonesian Inter-University Transport Studies Forum (IIUTSF), December 1998, Bandung (in Indonesian). (1998).
5. AWA Plessey. Bandung Area Traffic Control System: After Traffic Study, Bandung. (1997).
6. Low, D.E. A new approach to transportation systems modelling. “Traffic Quarterly” 26(3): pp. 391-404. (1972).
7. Nguyen, S. An algorithm for the traffic assignment problem. “Transportation Science” 8(3): pp. 203-216. (1974).
8. Willumsen, L.G. An Entropy Maximising Model for estimating trip matrices from traffic counts. Ph.D. Dissertation, University of Leeds. (1981)
9. Van Vliet, D. and Dow, P.D. Capacity restrained road assignment. “Traffic Engineering and Control” 20(6): pp. 226-235. (1979).
10. Tamin, O.Z. and Willumsen, L.G. Freight Demand Model estimation from traffic counts. Proceedings of the 16th PTRC Summer Annual Conference, Bath (UK). (1988).
11. Tamin, O.Z. Public transport demand estimation by calibrating a trip distribution-modal choice from traffic counts. Proceedings of the 5th CODATU Conference, Sao Paulo (Brazil). (1990).
12. Tamin, O.Z. Application of transport demand models for inter-regional vehicle movements in East Java. Proceedings of the 6th World Conference on Transport Research, Lyon (France). (1992).
13. Tamin, O.Z. and Soegondo, T. Freight demand modelling from traffic counts: A case study in Bali. Proceedings of the 5th World Conference on Transport Research, Yokohama (Japan). (1989).
14. Batty, M. Urban Modelling: Algorithm, Calibrations and Predictions. Cambridge University Press, Cambridge. (1976).
15. Wilson, A.G. and Bennet, R.J. Mathematical Methods in Human Geography and Planning. John Wiley and Sons Ltd. (1985)
16. Wilson, A.G. Entropy in Urban and Region Modeling. Pion Ltd, London. (1970).
17. DLGT. Bandung Area Traffic Control System, Final System Design, AWA Traffic System. (1996).
18. Tamin, O.Z., Soedirdjo, T.L., and Suyono, R.S. The impact of location and number of traffic counts in the accuracy of OD matrices estimated from traffic counts: A case study in Bandung (Indonesia), to be published in the 4th Journal of the Eastern Asia Society for Transportation Studies (EASTS). (2001).

ACKNOWLEDGEMENT

The authors would like to thank the University Research for Graduate Education (URGE) Project, Directorate General of Higher Education, Department of National Education for financially supporting this research ‘Graduate Team Research Grant, Batch IV, 1998/1999’ with the title: Dynamic Origin-Destination (OD) Matrices Estimation From Real Time Traffic Count Information under contract No. 029/HTTP-IV/URGE/1999.