A GRADIENT METHOD FOR THE ESTIMATION OF TRAVEL DEMAND USING TRAFFIC COUNTS ON THE LARGE SCALE NETWORK

Jonghyung KIM
Researcher
Department of Transport Policy
Korea Transport Institute
Tel:+31-910-3514
Tel:+31-910-3514
Email: Knight@koti..re.kr

Seungjae Lee
Assistant Professor
Department of Transportation Engineering
The University of Seoul
Dongdaemoon-Ku, Seoul, Korea
Tel: +82-2-2210-2172
Fax: +82-2-2215-5097
Email: Sjlee@uoscc.uos.ac.kr

Abstract: In this study, the surveyed TLFD(Trip Length Frequency Distribution) is determined as a criterion for the reliability of evaluating the true O/D matrix. The surveyed TLFD can be used to check the similarity between the surveyed (true) Trip Length Distribution and the Trip Length Distribution of the estimated O/D matrix by the traffic counted models. When the surveyed TLFD is similar to the estimated TLFD, the reliability and correctness of the estimated O/D are high. Therefore, the objective of this paper is the development of the travel demand (O/D matrix) estimation using traffic counts on the large-scaled network. The Gradient Method is used for the model and the multi-class assignment technique is used for the equilibrium loading procedure in the model. As a result, we have concluded that the multi-class based model has a closer value to the surveyed(true) TLFD than the TLFD of the estimated O/D matrix by the single-class based gradient method.

Key Words: Origin and Destination Matrix, Traffic Counts, Gradient Method, Multi-class Assignment.

1. INTRODUCTION

The origin and destination (O-D) matrix is one of the most important elements in transportation planning process. The accuracy of the O-D matrix plays key roles in the transport planning process in order to make and evaluate various transport policies. However, due to the nature of the O-D matrix, which is the desired people's or freight's movements on urban and regional space, it is very difficult and costly to estimate the O-D matrix. Traditionally, transport planners survey the O-D movements in order to estimate the O-D matrix. Even though the cost of the O-D survey requires high amounts of resources, the accuracy is relatively low. Even more, in the developing country, the transportation situation has changed very quickly and thus the transportation environment has been unstable. So the transportation planning should be frequently rectified according to the newly planned environment in order to capture the changed situation in the limited cost and time.

Until now, though most of the studies related to demand estimation method using traffic counts use methods based on single-class, travel demands or flows are made by mixing various classes in real networks.

In general, existing demand estimation methods based on traffic counts estimate O/D by converting a multi-class O/D matrix and traffic counts into a single-class O/D matrix and traffic counts through PCE conversion, and analyze a O/D matrix by dividing into a multi-class O/D matrix and traffic counts after multiplying an estimated O/D matrix by the fixed ratio of a single-class O/D matrix and traffic counts before PCE conversion. However, the merits of a demand estimation method based on multi-class calculate each route choice ratio about multi-class O/D, and maximize the estimation capability of multi-class by calculating each gradient, the reduction direction of objective function. Therefore, this study aims to establish a demand estimation method which considers congestion between class and class by using multi-class instead of single-class.

The main analysis direction of demand estimation methods based on traffic counts is to analyze the accuracy of estimation by using the method of error analysis. This has been conducted after comparing a basic O/D matrix with an estimated O/D matrix from a basically assumed demand in small-scaled and middle-scaled networks (see Lee and Kim, 2000). Such approaches are mainly used when we grasp the merits and demerits of models through observing the proposed models assuming true O/D matrix or true traffic counts which cannot be known in reality. However, though we are able to use such assumed network or true O/D investigating the application of models, it is nearly impossible to evaluate the reliability of estimated O/D in the case of real situations (large-scaled network). To solve these problems, we evaluated the reliability of the O/D matrix, which is estimated in a large-scaled network using TLFD (Trip Length Frequency Distribution). This is made by applying a multi-class based method which is proposed by this study to real data based on the personal trip survey of Seoul in 1996.

The result analyzed by multi-class based method has been similar to a surveyed TLFD, a true TLFD, in the TLFD analysis by zone. And, as a result of analyzing an observed traffic and an assigned traffic using chi-square method, it has been analyzed that the result of the multi-class based method has been superior to the result of the single-class based method. Therefore, the reason that the multi-class based method shows better results than the single-class based method is that the multi-class based method calculates congestion between vehicles and the rate of route choice by kinds of vehicles.

2. EXISTING STUDIES

Since the demand data cannot be observed directly, it must be collected by carrying out elaborate and expensive surveys, involving home or road based interviews or complicated number plate tagging schemes. By contrast, observed link volumes can be obtained easily either manually or automatically using mechanical or inductive counting devices. Thus, a considerable amount of research has been carried out to investigate the possibility of estimating or improving an origin-destination demand matrix with observed volumes on the links of the considered network. Many models have been proposed in the past such as Van Zuylen and Willumsen (1980), Van Vliet and Willumsen (1981), Nguyen (1982), Van Zuylen and Branston (1982) and Spiess (1987) among others. Most of these traditional approaches can be formulated as convex optimization problems in which the objective function corresponds to some distance function between a priori demand matrix and the resulting demand. The constraints are then used to force the assigned volumes to correspond to the observed volumes on the count post links.

Spiess(1990) proposed the gradient approach in order to estimate the O-D matrix using traffic counts. It is formulated as an optimization problem so as to minimize a measure of distance between observed and assigned volumes. The simplest function of this type is the square sum of the differences, which leads to the convex minimization problem. This is subject to the pseudo-function used to indicate the volumes resulting from an assignment of the demand matrix. The particular assignment model used must correspond to a convex optimization problem, in order for the objective function to be convex. Since the matrix estimation problem as formulated in the Spiess(1990) is highly underdetermined, it usually admits an infinite number of optimal solutions, i.e. possible demand matrices which all reflect the observed volumes equally well. In the actual planning context, we expect the resulting matrix to resemble as closely as possible the initial matrix, since it contains important structural information on the origin-destination movements. Therefore, just finding one solution to the problem in Spiess(1990) is clearly not enough.

If we would have a solution algorithm which inherently finds a solution close to the starting point, we could leave the objective function as is. Fortunately, the gradient method, also called the method of steepest descent, has exactly this property that we look for. It follows always the direction of the largest yield with respect to minimizing the objective function and, thus, it also assures us not to deviate from the starting solution more than necessary. In order to implement the gradient method, we also need to provide values for the step lengths. Choosing very small values for the step length has the advantage of following more precisely the gradient path, but has the disadvantage of requiring more steps. On the other hand, when choosing too large values for the step length, the objective function can actually increase and the convergence of the algorithm would be lost. Thus, the optimal step length at a given demand can be found by solving the one-dimensional subproblem.

3. MULTI-CLASS TRAFFIC COUNTS BASED GRADIENT MODEL

In this paper, the gradient model using multi class assignment is proposed. It is also formulated as an optimization problem like the single class assignment model. The objective function to be minimized is the square sum of the differences, which leads to the following convex minimization problem.

$$\min(t^m) = \frac{1}{2} \sum_{m \in M} \sum_{a \in A} (v_a^m - \overline{v}_a^m)^2 \tag{1}$$

subject to

$$v_a^m \geq 0, \quad \forall a, m$$

$$c(\mathbf{v}^*)(\mathbf{v} - \mathbf{v}^*) \geq 0, \quad \forall \mathbf{v} \in \mathbf{V}$$

subject to

$$\sum_k f_{kw}^m = t_w^m, \quad \forall w, m$$

$$f_{kw}^m \geq 0, \quad \forall m, w, k$$

$$v_a = \sum_m \sum_w \sum_k f_{kw}^m \delta_{ak}^w, \quad \forall a$$

where,

$$\mathbf{v} = \sum_m v_a^m b_a^m$$

$$\mathbf{v}^* = \sum_m v_a^{*m} b_a^{*m}$$

$$\delta_{ka}^w = \begin{cases} 0 \text{ if } a \notin k \\ 1 \text{ if } a \in k \end{cases}$$

b_a^m : congestion weight of each class

m: class

t^m : matrix of each class

v_a^m : assigned trip of class m in link a

$\bar{v}_a^m$: observed trip of class m in link a

$c(\mathbf{v})$: cost function

4. NUMERICAL EXAMPLE

The comparison of the multi-class and the single-class gradient model has been conducted using Seoul City network, which has 1,020 zones, 6,357 nodes (including 1,020 centroids), 19,127 links (including dummy links). Using two models, we compare results based on cordonline and screenline data as well as observed traffic counts. The gradient method has been known that it has made the least changed base O-D matrix. This means that the base O-D has been made in the base of the surveyed sample O-D and thus it has the surveyed TLFD information. If the base O-D has been changed little, this means that the estimated O-D using the Gradient Method has much information on the true people's or freight's movements. This leads to the good guideline to the usage of the traffic count based O-D estimation in practice and gives a confidence to the transport planner. It is because the traffic counted O-D estimation models give multiple solutions by its characteristics. Like below figure 1, the surveyed TLFD is compared with each TLFD of the estimated O-D. It shows that the most similar distribution of the surveyed TLFD is from the multi class based model.

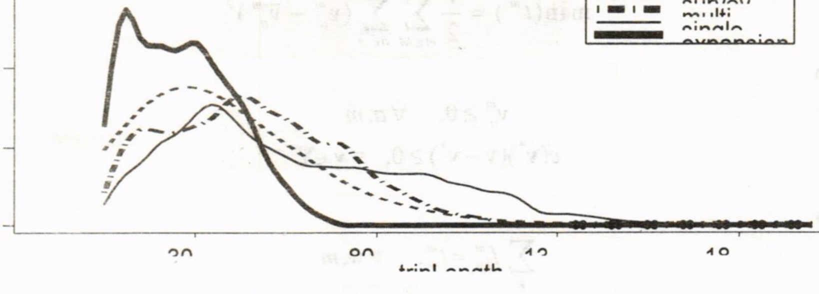

Figure 1. Comparison of the Estimated TLFDs and the Surveyed TLFD

The calculated and the surveyed TLFD should exhibit the following two characteristics: (1) the shape and position of both curve should be relatively closed to one another when compared visually, and (2) the difference between the average trip lengths should be within 3%.

According to two criteria above, the multi class based TLFD can only be an accepted solution. Except this multi class based TLFD, the difference of average trip length between the surveyed and the estimated does not meet the mentioned criteria.

Table 1. Comparison of trip lengths

	Average Travel Length (min)				Difference between average trip length (%)		
	① Survey	② Expansion	③ Single	④ Multi	Expansion (② - ①)/①	Single (③ - ①)/①	Multi (④ - ①)/①
Seoul	44.7	23.7	57.9	43.4	-47.0%	29.6%	-2.9%

The chi-square test is implemented to test homogeneity of the result by comparison of the surveyed TLFD with the others. As you can see in the Table 2 below, statistically, the most significant TLFD is multi-class based TLFD.

Table 2. Comparison of the model by a Chi-Square Method

	Multi	Expansion	Single
Level of Significance	0.05		
Statistics (Chi-square)	519.2068	3245.449	1090.537
Critical value	632.9418		
Result	Accept	Reject	Reject

First, the closest curve with the surveyed TLFD among three models is the multi class based TLFD. Second, the multi class based model shows the least difference between average estimated trip lengths and surveyed trip lengths. Finally, the chi-square test shows that the multi class based model has the most significant result statistically.

As a result, the multi class based model is a satisfactory model in the aspects of three criteria:(1)shape and position of curve, (2)difference of average trip lengths between surveyed and estimated, (3)chi-square test.

6. CONCLUSION

In this paper, the travel demand (O/D matrix) estimation model using traffic counts on the large-scaled network has been developed. The Gradient Method is used for the model and the multi-class assignment technique is used for the equilibrium loading procedure in the model. The Gradient method has been known that it has made the least changed base O/D matrix. This means that the base O/D has been made in the base of the surveyed sample O/D and thus it has the surveyed TLFD information. If the base O/D has been changed little, this means that the estimated O/D using the Gradient Method has much information on the true people's or freight's movements. This leads to the good guideline to the usage of the traffic count based O/D estimation in practice and gives a confidence to the transport planner. It is because the traffic counted O/D estimation models give multiple solutions by its characteristics.

In this study, the surveyed TLFD(Trip Length Frequency Distribution) is determined as a criterion for the reliability of evaluating the true O/D matrix. The surveyed TLFD can be used to check the similarity between the surveyed (true) Trip Length Distribution and the Trip Length Distribution of the estimated O/D matrix by the traffic counted models. When the surveyed TLFD is similar to the estimated TLFD, the reliability and correctness of the estimated O/D are high. Therefore, in this study, we analyzed the merits and demerits in each of a single-class based model and a multi-class based model in a large scale network. As a result, we have concluded that the multi-class based model has a closer value to the surveyed(true) TLFD than the TLFD of the estimated O/D matrix by the single-class based gradient method. Advantages of using the multiclass traffic counts based model are that the planners can estimate the O/D matrices for each class simultaneously and consider congestion effect between classes.

REFERENCES

a) Journal papers

Spiess H. (1987). A maximum likelihood model for estimating origin-destination matrices. **Transpn. Res. B 21,** 395-412.

Spiess H. and Florian M. (1989). Optimal strategies: A new assignment model for transit networks **Transpn. Res. B 23**, 83-102.

Van Zuylen H.J. and Branston D.M. (1982). Consistent link flow estimation from counts. **Transpn. Res. B 16**, 473-476.

Van Zuylen H.J. and Willumsen L.G. (1980). The most likely trip matrix estimated from traffic counts. **Transpn. Res. B 14**, 281-293.

Seungjae Lee and JongHyung Kim(2000), Comparative study on the O/D estimation using Gradient method and Generalized Least Square method, **Korea Society of Transportation**

Seungjae Lee, EuiYoung Son, JongHyung Kim(2000), Evaluation on the traffic count based O/D matrix using Trip Length Frequency Distribution, **Korea Society of Transportation**

JongHyung Kim, Seungjae Lee(2001), Development of a demand estimation method by using multiclass traffic assignment based on traffic counts, **Korea Society of Transportation**

b) Papers presented to conference

Willumsen L.G. (1984). Estimating time-dependent trip matrices from traffic counts. **Ninth International Symposium on Transportation and Traffic Theory**, VNU Science Press, 397-411.

Van Vliet D. and Willumsen L.G. (1981). Validation of the ME2 model for estimating trip matrices from traffic counts **Proceedings of The Eights International Symposium on Transportation and Traffic Theory**, June 1981.

Nguyen S. (1982). Estimating origin-destination matrices from observed volumes. **Proceedings of the First Course on Transportation Planning Models of the International School of Transportation Planning**, Amalfi, Italy, October 1982.

c) Other documents

JongHyung Kim(2000), Travel Demand Estimation from traffic counts using the Gradient Method – In the Focus of Multiclass traffic assignment on the Large scale network, PhD thesis, Univ. of Seoul

INRO Consultants Inc. (1989), **EMME/2** User's Manual.