DIGITALCOMMONS
@WAYNESTATE

Journal of Transportation Management

Volume 22 | Issue 1 Article 7

4-1-2011

Identifying traffic count posts for origin-destination matrix adjustments: An approach to actual size networks

Saeed Asadi Bagolee
Parsons, saeed.Asadi@parsons.com

Mohsen Asadi
Tarbiat Moallem University, mohsen.asadi@tmu.ac.ir

Lorna Richardson
Parsons, lorna.richardson@parsons.com

Recommended Citation

Bagolee, Saeed Asadi, Asadi, Mohsen, & Richardson, Lorna. (2011). Identifying traffic count posts for origin-destination matrix adjustments: An approach to actual size networks. Journal of Transportation Management, 22(1), 79-88. doi: 10.22237/jotm/1301616360

IDENTIFYING TRAFFIC COUNT POSTS FOR ORIGIN-DESTINATION MATRIX ADJUSTMENTS: AN APPROACH TO ACTUAL SIZE NETWORKS

Saeed Asadi Bagloee
Parsons

Mohsen Asadi
Tarbiat Moallem University

Lorna Richardson
Parsons

ABSTRACT

For transportation planners, the use of Origin-Destination (OD) matrix adjustment, is receiving considerable attention. However, there are concerns about the validity of results, primarily related to the number and location of traffic count posts. This leads to the question "What would be the best set of traffic count posts to use in OD matrix adjustment modules?" It has been proved that solving this problem is cumbersome. There have been several attempts (either exact or heuristic approaches) to address this problem. But due to the inherent complexities, there is no efficient and easy-to-use methodology able to address situations on the scale of actual cases. This study demonstrates a simple way of identifying traffic count posts tailored to deal with real-size cases. The proposed methodology is based on a maximum matrix coverage criterion. Using a limited number of incremental trials, a set of links whose traffic flows give maximum coverage of the demand and maximum fitness to the corresponding traffic count rates are identified as traffic count posts. The results show that more traffic count posts do not necessarily yield a better result. This article reports on a project conducted for the public works ministry of the UAE city of Sharjah.

INTRODUCTION

In transportation problems, the use of Origin-Destination Matrix Adjustment (OD-MA) based on traffic count data is receiving considerable attention from practitioners. This is due to the fact that the approach provides a cost efficient alternative to the time consuming and expensive traffic surveys required to develop OD matrices. In addition, after a couple decades of research, most transportation planning software provides this application. However, the extensive utilization of the OD-MA has faced some obstacles largely dealing with the number and location of the traffic count posts.

There are various methodologies which provide solutions to the OD-MA problem such as Spiess (Spiess, 1990; Nguyen, 1984; Cascetta, 1984; and Yang, 1994). Most of these approaches formulate a convex optimization problem in which some sort of distance function $Z(\hat{D}, D)$ between an initial demand matrix $\hat{D}$ and adjusted demand D is developed. In order to achieve assigned volumes v_a relatively close to observed volumes $\bar{v}_a$ on the count posts (links) $a \in \bar{A} \subset A$ (A is set of network's link), some constraints are embedded in the formulation. In all, the primary input after the initial matrix is the traffic counts. Intuitively, the set of traffic counts must observe some considerations such as:

- The traffic count rates must be consistent.
- Traffic count posts must be independent.

- In general, traffic count posts must represent as much travel demand as possible. Yang et. al. *(1998)* has defined this consideration with three rules.

- Apart from the above items, from a technical perspective, there are some other considerations such as: (a) the count posts should not capture a lot of intra-zonal trips since these trips will not be accounted for in the traffic assignment. (b) The count posts should not be placed close to zone connectors, because, to achieve a better fitness to the traffic count rates, the corresponding zones would be biased to observed volumes of the corresponding count posts.

Due to the nature of the OD-MA procedure (simply in terms of the unbalanced number of unknown variables and equations); the outcome solution may not be unique. This fact, together with the considerations listed above, has raised a substantial concern about possible perturbations consequently being imposed on the initial matrix (in terms of trip distribution pattern, total number of trips, etc.).

The initial matrix is typically developed from an elaborate and expensive survey (such as home or road side interviews ...) which contains substantial structural information on the origin-destination movements. Therefore the final adjusted solution (out of so many solutions) must not vary significantly from the initial matrix. This is a very strong criterion in which no compromise is tolerable. There have been some studies addressing the uniqueness of solution by introducing more constraints and criteria or a secondary objective function to select the most desirable solutions. Yang et. al. *(1998)* and Chootinan et. al. (2005) set up some rules, such as an OD covering rule, maximal flow fraction rule, maximal flow-intercepting rule and link independence rule, and proposed integer linear programming models *(*Yang et al, 1998*)* or a bi-objective problem (Chootinana, Chena, and Yang, 2005). They developed heuristic solution methods to determine the counting links satisfying the established rules. Their methodologies were not, however, tested by a real case study. LeBlanc et. al. *(1982)* proposed a partial Lagrangian method to choose the nearest solution (OD matrix) to the initial matrix among all feasible solutions. Computational results from the application to a small network in Sioux Falls, South Dakota with 76 links were presented. Spiess *(1990)* made a great effort by introducing a relative version of gradient method in which the adjusted matrix would be proportional to the initial matrix so as not to deviate dramatically from the initial matrix.

It is worth noting that, in contrast to Spiess approach, most of the developed methodologies (Nielsen, 1998); Ortuzar and Willumsen; 1990; and Willumsen, 1981) and commercial planning software applications (TransCAD, 1996) yield an adjusted non-zero matrix with a good fitness to the traffic count rates on the basis of a zero-out initial matrix. Also, the implication of Yang et. al.'s *(1994) work,* wherein the OD-MA problem can be greatly simplified under certain conditions, shows that the OD-MA applications are very fragile. This may result in many good solutions being discarded. In this regard, the importance of adopting a proper OD-MA module associated with proper traffic count posts deserves more attention so as not to deteriorate the initial matrix.

There may be various interpretations of the traffic count post problem. For instance, given that conducting traffic count surveys is not free of charge (and budget always is limited), one may want to know the location of the minimum number of link count posts in order to determine the traffic volume of the entire network. This problem in math and computer science is called a Sensor Location Problem or Dominating Path Problem. To provide a sense of the complexity of these kind of problems, Bianco et. al. *(2006)* proved that the problem is in the complexity order of NP-complete.

In this study, without getting overly absorbed in the problem's complexity, a practical version of the problem is addressed as follows: "There is a set of traffic count rates produced from junction and corridor analysis (as part of regular activities in a Traffic Impact Study -TIS project); but what is the best subset to feed into the OD-MA module in order to have a reliable adjusted model?" The members of the traffic count rates are henceforth referred to as "Candidate Traffic Count Posts" (or CTCP).

This study presents an approach to deal with real size cases using an actual project conducted for the UAE's public works ministry. First CTCPs are prioritized and sorted according to their demand coverage. Second, through an iterative and incremental process, starting from the top prioritized CTCPs, a subset of CTCPs is "fed" into the OD-MA module. Spiess' algorithm (via demadj.mac; macro feature of EMME3 (Spiess, 1990)) based on least error between counts and volumes is then engaged to carry out OD-MA. Next, the fitness (R^2-index) of the assignment volumes to the corresponding CTCPs would be a key parameter to decide which subset of the CTCPs must be chosen as the "traffic count posts." Application to the case study showed there was an optimum number of a CTCPs with maximum R^2-index (i.e. feeding the OD-MA module with more CTCPs does not necessarily yield better result).

METHODOLOGY

From a practical perspective, given a traffic network and a set of traffic count rates (usually collected during TIS projects) adjusting outdated OD matrixes to the traffic counts is desirable. Practitioners' and researchers' experiences reveal that feeding the OD-MA module with all the counts might have adverse effects by deteriorating the number and distribution of trips. Thus, in simple language this question arises: "Given a set of traffic count rates - should all the counts serve as inputs for the module? If not, which count rates should be used?" This study answers the question for a real size case.

In order to select a subset of CTCPs, Yang et. al. (1998) proposed some rules that they derived from empirical observations and common sense as follows:

- Rule-1: The OD Covering rule - some fraction of the trip for each OD pair must be covered.
- Rule-2: The Maximal Flow Fraction rule - for a given OD pair, the count post should be identified in a way that, the largest fraction of flow for that OD pair is obtained.
- Rule-3: The Maximal Flow Intercept rule - given a set of candidate posts, choose the ones that have the greatest number of OD pairs traversing them.

In principle, these rules are all good, however, in practice; rules 2 and 3 often come into conflict with each other. In addition, as discussed before with respect to the complexity of the problem, the proposed solution methodologies are not able to tackle real size cases (Yang and Zhou, 1998; Chootinana, Chena and Yang, 2005). Since the primary objective of this study is to ensure its applicability to the real world, even if this involves compromising some purely mathematical aspects of the problem, this study adopted "more matrix demand coverage" (which can be interpreted as a general aggregation of the triple rules) as a benchmark to prioritize and then select the best collection of count posts.

Our approach then is carried out as follows. Initially, the original (initial) demand matrix is assigned to the network so that the traffic volumes of all the links are saved. Also, travel time emanating from assigning the initial matrix on the network is saved and the times on all the links of the network are preserved. Then a candidate count post with maximum traffic volumes is labeled as the first prioritized candidate posts. In order to find the next one, the previous prioritized candidate post is removed from the candidate post set. Thus the part of the demand matrix corresponding to the prioritized candidate is removed from the matrix as well. The consequent matrix could be called a "truncated matrix". The truncated matrix is assigned on the network while the travel time has been preserved as it was for the initial assignment. By doing so, the resulting traffic flow simply is the original traffic flow minus the flow corresponding to the previous prioritized post(s). Again a candidate post with maximum current flow is labeled as the next prioritized candidate posts. This process may be repeated until a sorted and prioritized set of CTCPs is identified.

This heuristic approach to addressing the problem has some significant advantages. First, executing this concept even for practitioners is very easy. Commercial planning software provides useful procedures called "Select Link Analysis" in which an OD matrix corresponding to desirable links can be distinguished from the original matrix. Furthermore, Emme3 provides an easy way to conduct the prioritization process through a macro called cntposts.mac (INRO Consultant, Inc., 2010) in which additional options of auto assignment are used. The user simply enters the initial matrix and the set of candidate posts. Within a very efficient computing time, the macro computes and tags the amount of demand coverage for the candidate posts (i.e. more coverage means higher priority).

Secondly, if the initial matrix and the network are "reliable," the results would respect all of Yang's rules in one way or another (reliability taken here to mean observing consideration-4 presented earlier). Third, the magnitude of the last traffic flow on each count post is an indication of how important the count posts (prioritization of importance) are. This property can be important since some algorithms, such as Spiess, are able to accept some sort of the weights for count rates. Thus the adjusted matrix would be biased to those count rates with more weights. For instance one may want to have the adjusted matrix more closely reflect count rates along highways and expressways rather than local and access roads. By having the set of traffic count posts, the OD-MA module is executed.

Spiess' methodology based on the gradient method to minimize distance between counts and assigned volumes as a convex minimization problem is:

$$\min_{D} \; Z(\hat{D},D) = \frac{1}{2} * \sum_{a\in A} (v_a - \bar{v}_a)^2 \qquad (1)$$

Subject to

$$v_a = assign(D) \qquad for \;\; \forall a \in A \qquad (2)$$

Wherein $assign(D)$ indicates that v_as are the volumes emanating from a traffic assignment in which an equilibrium traffic flow results. Since the expressed problem is highly undetermined, an infinite number of solutions (all yielding a close fit to the observed volumes) are expected. Due to the substantial structural information of the initial matrix; the proximity of the solution to the initial matrix must be noted. Thus Spiess has proposed a transformed gradient method to solve the problem (1-2) in which the gradient is based on the relative change to the demand as follows:

$$D_i^{\ell+1} = \begin{cases} \hat{D}_i & for \;\; \ell = 0 \\ D_i^{\ell}\left(1 - \lambda^{\ell}\left[\dfrac{\partial Z(\hat{D},D)}{\partial D_i}\right]_{D_i^{\ell}}\right) & for \;\; \ell = 1,2,3,\ldots \end{cases} \qquad (3)$$

Where λ^{ℓ} is the size of the move along the steepest descent $\left[\frac{\partial Z(\hat{D},D)}{\partial D_i}\right] D_i^{\ell}$

at iteration ℓ. By using relative gradients, the solution algorithm becomes multiplicative in initial demand $\hat{D}_i$ so a change in demand is proportional to the initial demand. This module has been implemented in Emme3 and is called demadj.mac (Spiess, 1990). Finally, the R^2-index between all the survey counts, including those fed into the module, versus the assigned volumes is used as a measure to identify which set of CTCPs gives the better result.

Out of all the initial survey count posts, 10% are prioritized by executing cntpost.mac and designated to be fed into demadj.mac. Through an incremental process, for each next attempt a further 5% are added to the already fed CTCPs. This process continues until all the survey is taken as count posts. In the end, the attempt with the maximum R^2-index, along with some other considerations, can be chosen as the updated model.

UAE CASE STUDY

A model of the city of Sharjah, UAE comprises 481 zones, 10,426 nodes and 26,294 links. There are total trips of 182,128 and 182,908 for the AM and PM peak hours respectively. A traffic survey which was carried out over 18 junctions plus 8 roads accumulated up to 281 movements. Figure 1 depicts the traffic survey locations. The algorithm ran 9 times starting with 10% of all traffic surveyed (28 candidate posts) and then up to 50% (at which no more improvement in the R^2-index was observed).

The Spiess module provides the facility to weight specific count posts in order to attain a more desirable pattern. For instance, one may want to get a conservative pattern in which the results guarantee higher rates of traffic counts. Thus a logistic function varying from 1.00 for low rates up to 2.0 for the highest volume is adopted here as follows:

$$link'sWeight = 2.\frac{100}{100+100*e^{5.\frac{traffic\ count\ rate}{\max imum\ observed\ traffic\ count\ rate}}}$$

(4)

In every attempt, the Spiess module is set for 8 iterations so as to guarantee fitness of $R^2 = 0.95$ or above for the (only) fed counts and corresponding assigned volumes. At the end of each attempt the R^2-index for all the 281 CTCPs (including non-fed and fed count posts) against the corresponding assigned volumes is calculated. Table 1 indicates results of the incremental tries.

In Table 1 the introduced indices are:

- An R^2-index: an index for overall performance of the methodology

- Total travel demand and average travel time: shows how the adjusted matrix differs from the initial matrix are.

Figure 3 depicts the changes of the listed above indices over incremental numbers of traffic count posts.

Figure 3a clearly shows that there is an optimum collection of CTCPs to be utilized since feeding the algorithm with more posts only results in deterioration in the overall convergence of the algorithm. For both AM and PM; using 40% of CTCPs (112 count posts) has achieved around 80% overall fitness. Figures 3b and 3c indicate that in terms of closeness to the initial matrix for a low number of traffic count posts the algorithm behaves chaotically and is not reliable. As the number of count posts increases the results assume a monotone shape. Provided the initial matrix is accepted, an adjusted matrix close to the initial matrix in terms of average travel time and total amount of trips may be taken. For

FIGURE 1
SHARJAH, UAE TRAFFIC SURVEY LOCATIONS

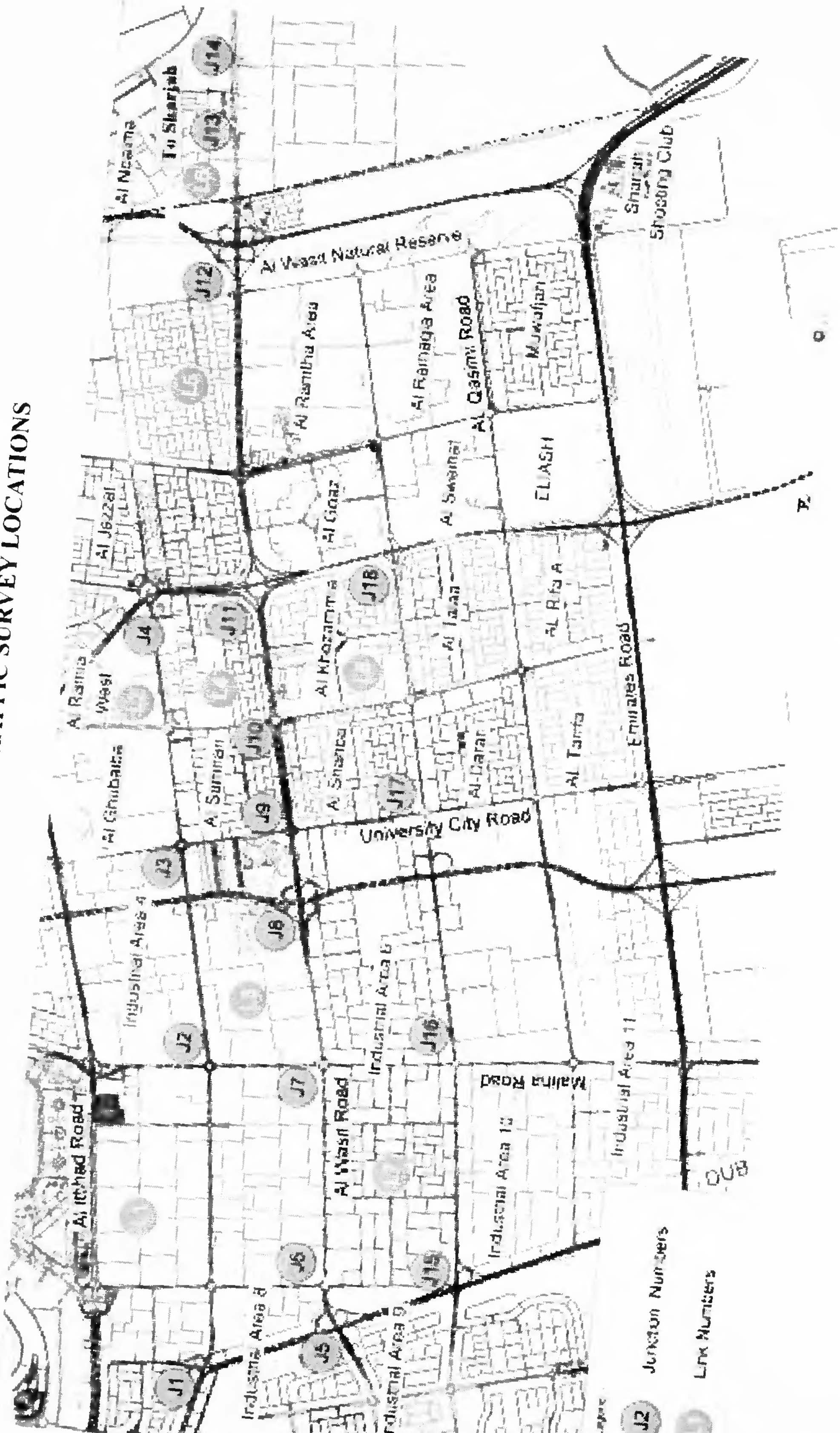


FIGURE 2
LINKS WEIGHT FACTORS

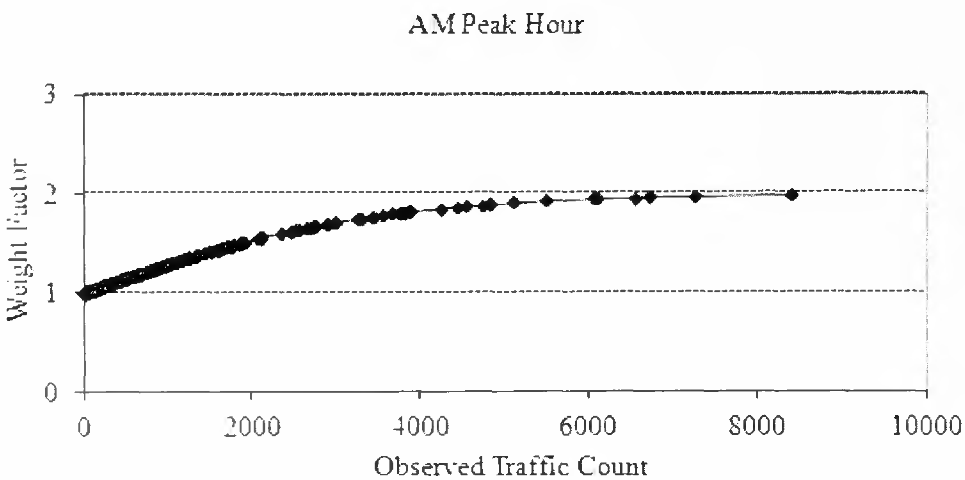
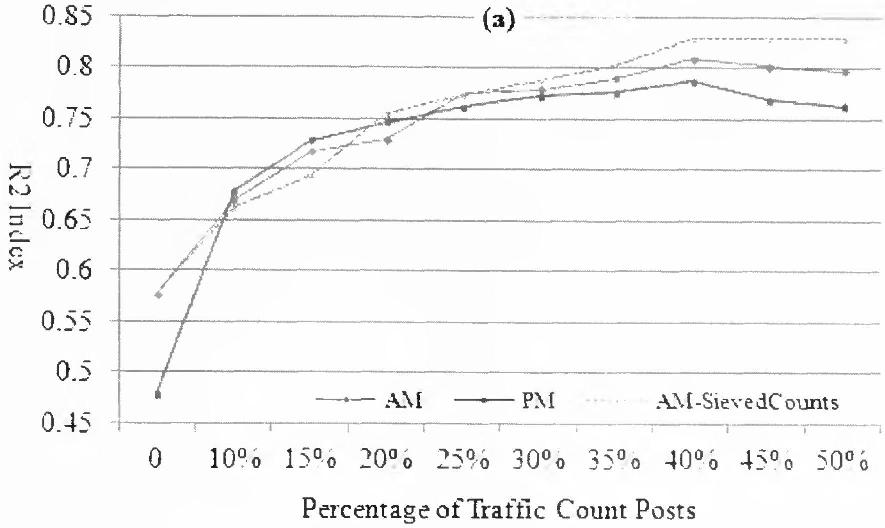
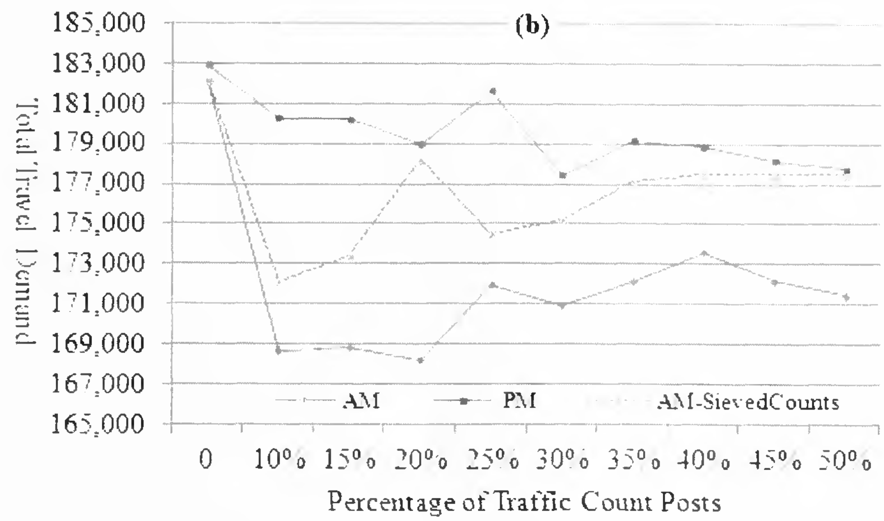


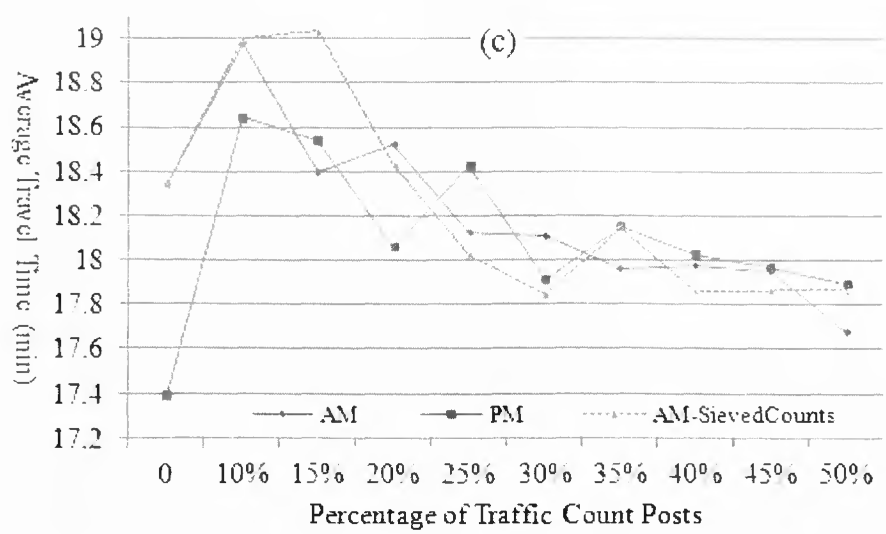
TABLE 1
MODEL'S RESULTS FOR INCREMENTAL PERCENTAGE OF TRAFFIC COUNT POSTS

			AM			PM		
Try No	Percentage of Fed Traffic Count Posts	Number of Fed Count Posts	R^2 Index	Total Travel Demand	Average Travel Time (minute)	R^2 Index	Total Travel Demand	Average Travel Time (minute)
0	0	0	0.5766	182,128	18.342	0.4783	182,908	17.392
1	10%	28	0.6690	168,695	18.981	0.6784	180,276	18.642
2	15%	42	0.7171	168,837	18.402	0.7284	180,228	18.537
3	20%	56	0.7294	168,194	18.519	0.7463	179,007	18.056
4	25%	70	0.7746	171,946	18.125	0.7618	181,609	18.420
5	30%	84	0.7787	170,979	18.106	0.7730	177,456	17.912
6	35%	98	0.7899	172,150	17.957	0.7760	179,570	18.147
7	40%	112	0.8086	173,612	17.980	0.7871	178,851	18.023
8	45%	126	0.8018	172,193	17.949	0.7689	178,138	17.968
9	50%	140	0.7976	171,468	17.676	0.7623	177,743	17.895

FIGURE 3
ALGORITHM RESPONSES OVER NUMBER OF FED COUNT POSTS







instance in the AM, the maximum R^2-index occurred at 40% of traffic count posts and at which point the total trip is at the nearest distance to the initial matrix.

Number of Count Posts

It is useful to consider why, counter-intuitively, more count posts do not necessarily provide better results (higher R^2). It is possible that beyond the optimum set of traffic count posts, the additional traffic count posts convey no additional information. Generally speaking this may be due to installing count posts at some linearly dependent locations with the optimum posts, survey errors or selecting unimportant locations such as seldom used local roads. In order to demonstrate that the poor count posts have adverse effects a new run is conducted on a selected set of counts posts rather than the initial set for the AM peak hour. We set up some thresholds to discard the poor posts from the initial set in order to have a selective set of count posts before launching the methodology. First we calculate the traffic survey rates-per capacity ratio (known as V/C in transportation literature) for all the candidate posts. To avoid major survey errors and low-profile local roads, the candidate posts with a V/C ratio greater than 20% or with traffic survey rates greater than 360 are used as the selective count posts set and the remainder discarded. This selective set simply is called AM-SievedCount which contains 161 count posts out of 286 initial count posts. The threshold of 20% for V/C is an arbitrary parameter embedded to exempt the methodology from fitting low profile count posts. Similarly the minimum traffic survey of 360 can be seen as passing at least 1 car every 10 seconds.

The methodology as described was run on AM-SievedCount. The result is shown in Figure 3. Figure 3(a) demonstrates that during successive tries the algorithm steadily rises to a saturated level (7st attempt) at which maximum (possible) fitness is achieved. Beyond this level no more candidate traffic posts belonging to AM-SievedCount would be selected due to linearly located count posts. Figure 3(b) demonstrates that adopting a sieved set of count posts may produce a reliable result in the sense of closeness to the initial matrix. The above discussion once again highlights the importance of properly and carefully identifying the traffic count posts. These results lead us to the point that OD-MA is not always predictable or straightforward and should therefore not be used as an alternative to standard procedures for developing trip tables. This situation is exacerbated further if great care has not been taken in identifying the count posts. In addition, and counter-intuitively, more count posts do not yield a better result.

CONCLUSION

This paper introduces an easy and efficient approach to the problem of selecting the best set of traffic count posts for the purpose of the OD Matrix Adjustment (OD-MA), applicable to real networks. From a traffic survey conducted for the Traffic Impact Studies (TIS) a set of traffic count rates (called candidate traffic count posts or CTCPs) is available. First, CTCPs are prioritized and sorted according to their demand coverage. Second, through an iterative and incremental process, starting from an initial number of top prioritized CTCPs (10%) and incremental rates (5%) up to an endpoint, a subset of CTCPs is designated and "fed" into the Spiess' OD-MA module. Then the fitness (R^2-index) of the assignment volumes to the corresponding CTCPs would be a key parameter to decide which subset of the CTCPs must be chosen as the "traffic count posts". Application to the case study showed there was an optimum number of CTCPs with a maximum R^2-index. Feeding the OD-MA module with more CTCPs does not necessarily yield better result).

During the case study an important observation was achieved: counter-intuitively, by feeding more traffic counts the module achieved a better fit, but overall, it deteriorated in the size and distribution of trips. For instance, in the case of Sharjah, UAE, 112 count posts (40% of all the traffic count) yields a maximum R^2-index for all the traffic count posts versus corresponding assigned volumes. This study has an important implication: before using the OD adjustment modules it is necessary to first identify which links should be taken as traffic count posts.

No matter how reliable the count posts are, even by accommodating relative versions of gradient methods so as to have an adjusted matrix close to the initial matrix; great care must be taken when the OD-MA application is used. A visible discrepancy between the final adjusted matrix at the highest overall fitness (maximum of R^2-index) and the initial matrix in terms of total trip rates and average travel time was observed.

Acknowledgement

The authors are greatly indebted to unanimous reviewers and the editors for their comments that contributed in the final version of the paper .The authors wish to thank Dr. Hossein Poorzahedy from Sharif University of Technology, Tehran, Iran and Smith Myung of Cambridge Systematic Inc. for their helpful comments in improving the paper.

REFRENCES

Bianco, L., G. Confessore, and M. Gentili (2006), "Combinatorial Aspects of the Sensor Location Problem," *Annals of Operations Research,* 144 (1): 201-234.

Cascetta, E. (1984), "Estimation of Trip Matrices from Traffic Counts and Survey Data: A Generalized Least Square Estimator," *Transportation Research*, 18B (4/5): 289-299.

Chootinana P., A. Chena, and H. Yang (1982), "A Bi-Objective Traffic Counting Location Problem for Origin-Destination Trip Table Estimation," *Transportmetrica,* 1 (1): 65-80.

INRO Consultant Inc., EMME Support Center-http://www.inro.ca/en/download/public/share/macros/cntpost.mac, Accessed July 28, 2010.

LeBlanc, L.J., and K. Farhangian (1982), "Selection of a Trip Table Which Reproduces Observed l Link Flows," *Transportation Research*, 16B (2) 2: 83-88.

Nielsen, O.A. (1998), *Chapter 13: Two New Methods for Estimating Trip Matrices from Traffic Counts*, Travel Behavior Research: Updating State of Play, Edited by: Juan de Dios and Others, Elsevier Publishers.

Nguyen, S. (1984), "Estimating Origin-Destination Matrices from Observed Flows," *Transportation Planning Models, F.,* Florian Eds., Elsevier, Amsterdam, The Netherlands: 363-380.

Ortuzar, J.D.D., and L.G. Willumsen (1990), *Modeling Transport,* Chichester, England: John Willey and Sons.

Spiess, H. (1990), *A Descent Based Approach for the OD Matrix Adjustment Problem*, Center for Research on Transportation, University of Montreal, Canada, CRT Pub. No 693.

TransCAD (1996), *User Manual, Version 3.0. Caliper Corporation,* Newton, Mass.

Willumsen, L.G. (1981), "Simplified Transport Models Based on Traffic Counts," *Transportation*, 10 (3): 257-278.

Yang, H., Y. Iida, and T. Sasaki (1994), "The Equilibrium-Based Origin-Destination Matrix 3 Estimation Problem," *Transportation Research*, 28B (1): 23-33.

Yang, H., and J. Zhou (1998), "Optimal Traffic Counting Locations for Origin-Destination. Matrix Estimation," *Transportation Research*, 32B (2): 109–126.

AUTHOR BIOGRAPHIES

Saeed Asadi Bagloee is with PARSONS in Dubai, UAE. Email: Saeed.Asadi@parsons.com.

Mohsen Asadi is with the Tarbiat Moallem University in Karaj, Iran. Email: mohsen.asadi@tmu.ac.ir

Lorna Richardson is with PARSONS in Dubai, UAE. Email: lorna.richardson@parsons.com