International Journal for Traffic and Transport Engineering, 2014, 4(4): 415 - 424

DOI: http://dx.doi.org/10.7708/ijtte.2014.4(4).05 *UDC: 629.11*

ESTIMATION OF THROUGH TRIPS USING EXISTING TRAFFIC COUNTS

Tahmina Khan[1], Michael Anderson[2]

[1, 2] *Civil and Environmental Engineering Department, University of Alabama in Huntsville, Huntsville, AL 35899, United States*

Received 22 April 2014; accepted 29 July 2014

Abstract: Through trips are a concerns as these trips contribute to roadway congestion that must be accommodated, but for which there are no simple methods to determine. Concerns for safety and a need to avoid inconveniencing drivers have limited the use of traditional vehicle surveys. Previous research developed equations to estimate through trips in communities, requiring data on street locations, traffic volumes, community demographics, economic data and geographic data. Alternatively, using cell phones or Bluetooth devices to collect traveler information have not always been accepted by the general public and are often seen as intrusions into driver privacy. Transportation professionals are in need of a methodology to estimate through trips to improve the transportation planning process and better allocate resources for roadway infrastructure investment. This research utilizes roadway connectivity and traffic count data to estimate through trip numbers. The methodology determines through trips using roadway counts collected as part of a routine traffic monitoring procedure. The methodology has been tested against a community that underwent a Bluetooth data collection study. The outcome of this research will benefit any community with a traffic monitoring program where through trip patterns could be used to improve resource allocation.

Keywords: through trips, external-external trips, origin destination matrix.

1. Introduction and Background

Through trips, or pass through trips, are great concerns for small and medium sized communities as these trips contribute to congestion on roadway infrastructure of the community that must be accommodated, but for which there are no simple methods to determine. The complete construction of a through trip table also involves determining the fraction of trips at an external station that are through trips, as opposed to trips having one end internal to the study area (Horowitz and Patel, 1999). Trips can be termed as through trip or external-external (E-E), where origin and destination of trip fall outside the community (Han, 2007). External surveys, conducted at external stations, obtain through trip information through license plate surveys, roadside handout surveys, roadside interview surveys, and roadside interview combined with handout surveys (Han, 2007). However, the use of surveys has diminished due to rising costs, traffic delays and safety issues. Attempts have been made to use cell phone records or Bluetooth capture devices to collect traveler information. Unfortunately, these have not always been accepted by the

[2] Corresponding author: andersmd@uah.edu

general public and are often seen as intrusions into driver privacy.

The main goal in estimating through trips is to predict the total number, or percent, of trips that would be passing through and the distribution of trips between external stations. To remove the issues associated with surveys, regression models or synthetic procedures have been developed. It should be noted that through trip results are not always transferable between areas as they are heavily depend on the specific location, size, roadway network and relationship to other communities.

The principle methodology for determining through-trip rates cited in recent literature involves the application of a series of regression models that were developed based on external station surveys. The models predict the external trip exchange based on highway functional classification, the average daily traffic (ADT) at the external station, the percentage of trucks (excluding vans and pickups), the percentage of vans and pickups, route continuity and the population of the study area. As an alternative, Anderson et al. (2006) presented a spatial economic model to synthesize a through trip table using surrounding communities and their impact, was shown to be more accurate than the common regression-based model for limited applications. Han (2007) updated through trip estimation procedures with new survey data and to account for geographic and economic explanatory factors that can be applied in any US small and medium urban areas. The drawback of the current through trip models is that, while applicable in small and medium sized communities, there is still an element of borrow from other studies and hope that the equations/methods will be transferable (Anderson et al., 2006).

Transportation professionals need a methodology to estimate through trips to improve the transportation planning process and better allocate resources for roadway infrastructure investment. This paper presents research conducted that utilizes roadway connectivity and traffic count data to estimate through trip patterns. The methodology developed determines travel demand between origin-destination locations from the actual roadway counts, collected during routine traffic monitoring procedures. The methodology has been tested against and performed similarly to a community that underwent a Bluetooth data collection study. The outcome of this research will benefit any community with a traffic monitoring program where through trip patterns could be used to improve resource allocation.

2. Model Development Methodology

The methodology used to generate through trip rates for a small and medium sized community is to develop an origin/destination (OD) matrix from existing traffic counts without requiring an existing OD matrix by highlighting entropy maximizing work by Wilson (1970). Traffic counts, viewed as the combination of a trip matrix and a route choice pattern, can give direct information about the sum of origin/destination pairs that use a particular roadway in the network (Ortuzar and Willumsen, 1994).

After identifying the roadways used by the trips from origin to destination, the flows for roadways are expressed mathematically in the following Eq. (1) (Ortuzar and Willumsen, 1994):

$$V_a = \sum_{ij} T_{ij}^a \, p_{ij}^a \ , 0 \leq p_{ij}^a \leq 1 \qquad (1)$$

Where the flow (V_a) for link a is the summation for all the trips on that link, p_{ij}^a is the proportion of trips (T_{ij}^a) from Zone i to Zone j travelling through the link. Now, the proportion variable is determined by the type of trip assignment technique which can be classified into two main ways, such as proportional and non-proportional assignment. In proportional assignment the proportion of drivers choosing each route are independent from the flow levels and calculate the probability of traffic coming from a particular origin and destination for each link flow. The most common use of the proportional assignment is through the use of the all-or-nothing assignment method, illustrated in the following Eq. (2), while some stochastic methods will give values that range from zero to one to variable p_{ij}^a (Ortuzar and Willumsen, 1994):

$$p_{ij}^a \begin{cases} 1 \text{ if trips from origin i to destination j use link a} \\ 0 \text{ otherwise} \end{cases} \qquad (2)$$

Under congested conditions, non-proportional assignment would allow for trips to take paths other than the shortest travel time and causes the proportion of travelers on each link to not depend on link flows. It is an iterative approach where a set of route choice proportions are assumed then a trip matrix is estimated, loaded onto the network and a new set of route choice proportions are calculated, and the process is repeated until the route choice proportions and trip matrix are similarly consistent. Therefore, the interdependency between the route choice proportions and the trip matrix is the main downfall to this technique (Ortuzar and Willumsen, 1994).

One way of implementing non-proportional assignment is through the use of incorporating equilibrium to the traffic flows that assigns link cost functions, link travel cost, and path travel cost to the network as a way to minimize travel costs (Xie et al., 2010; Bera and Rao, 2011). A target trip matrix is required to reproduce the observed traffic counts in the equilibrium approach (Xie et al., 2010; Abrahamsson, 1998). These models were developed on small test networks and their applicability on a large network is not ensured (Bera and Rao, 2011).

3. Case Study

To test the methodology presented for building an OD matrix to determine through trips, a case study was conducted using Brazos County, Texas network. The location was selected as there was a recent E-E study preformed that used Bluetooth data collected from cell phones, which was used as a validation of data set.

3.1. Selection of Study Area

Defining the territory of a study area depends on where the external zones/stations located that are the main generator or attractor of through trips. Based on the road network, boundary and location of external zones internal and external zones have to be identified and labeled by number on a printed map. Brazos's external zone ranges from 1 through 13 and internal zone ranges from 14 through 58 while intersected nodes start from 100 to 219.

3.2. Required Data for Analysis and Validation Process

The aforementioned road network map preferably needs to have the traffic count for all links and a defined uniform scale because if the distance or travel time between zones is not available then distance between nodes has to be measured manually from the map and tabulated properly. Also, the actual percent through trips and total trips of the external stations will be required for validation purposes. It can be noted that the year must be same for the traffic counts of the study area, the actual percent through trips and total trips of the external stations. District traffic data from the Texas Department of Transportation (TxDOT) supplied the necessary traffic counts and street network map for Brazos County (Divison, 2012). Fig. 1 shows the study area namely Brazos County with 13 external stations (ES) (Farnsworth et al., 2011). Table 1 presents 2011 Bluetooth output of 13 external stations where the Count column has been collected from TxDOT traffic count, was placed and aligned with Bluetooth output for comparison purposes (Farnsworth et al., 2011; Divison, 2012). It can be seen that there are discrepancies between Total trips and Count Columns; however, as the goal of this study was the external trip numbers and percent, the values from the Bluetooth study were used.

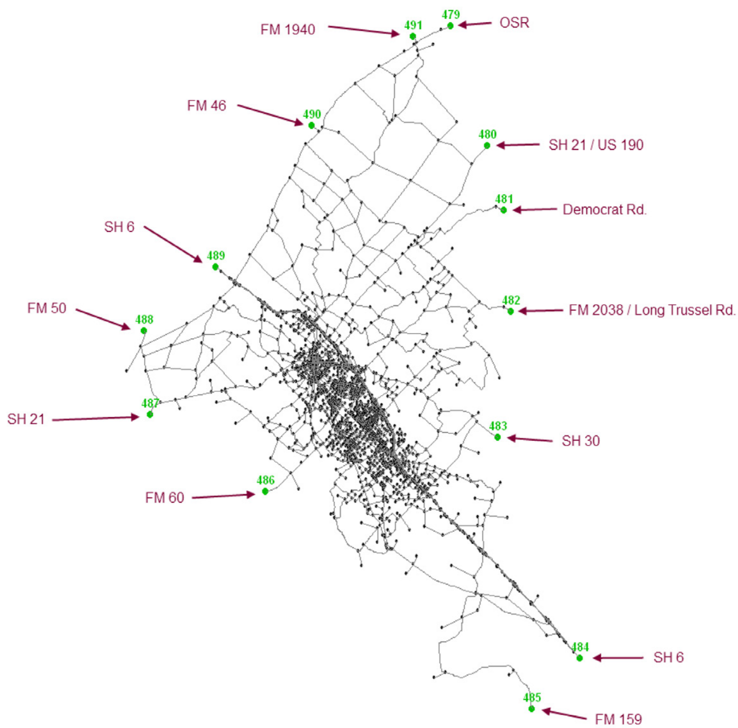

Fig. 1.
Brazos Study Area with 13 External Stations

Table 1
2011 Bluetooth Test Output

Station	ES No	Through Trips	Local Trips	% Percent Through	Total Trips	Count
479	1	91	1,028	8.1%	1,119	1,200
491	2	21	808	2.5%	829	1,050
490	3	65	1,925	3.3%	1,990	2,300
489	4	3,506	21,156	14.2%	24,662	19,000
488	5	109	871	11.1%	980	1,250
487	6	863	12,146	6.6%	13,009	12,400
486	7	216	8,474	2.5%	8,690	9,300
485	8	14	418	3.2%	432	450
484	9	3,405	22,476	13.2%	25,881	26,000
483	10	361	6,410	5.3%	6,771	6,000
482	11	0	207	0.0%	207	490
481	12	4	152	2.3%	156	700
480	13	538	6,812	7.3%	7,350	6,900

3.3. Determining Shortest Route between Zones

3.3.1. Preparing Input Database

Links, or roadway sections, with traffic count are represented using two adjoining points, NODE A and NODE B, as defined in our study. The nodes that we used to represent zone centroids were labeled differently than intersected nodes to ease in identification and post-processing. The corresponding link distance is recorded and used for routing purposes. Brazos study area was divided into 58 zones and access nodes were given where there is a change in traffic count. Access nodes and intersected nodes were numbered 100 through 219. Input file with three columns thus contains link, zone and distance information of a study area.

3.3.2. Building Shortest Route Algorithm

The movement from one specific zone to the other was coded in a MATLAB script file that developed tree diagrams to determine shortest path between origin and destination. The MATLAB script read the excel input file and starts finding route from zone 1 to other 58 zones such as 1 to 2, 1 to 3 and so on as incremented in a loop statement which is nested by another loop to find the same for zone 2, 3 and so on. Each route was represented as the connecting nodes between the origin and destination and total distance of corresponding links. The final output was formed to produce an array/matrix with all movements between zones in rows and the each roadway link presented in columns was attributed with a value of 1 if that roadway link was used on the shortest path between origin and destination or else a 0 if not used.

3.3.3. Running and Exporting Output Table

This algorithm can be run by feeding in the input database for any study area and desired 0 and 1 output table can be exported and pasted in excel spread sheet.

3.4. Developing OD Spread Sheets for Initialization and Running Iterations

The assignment was conducted through the use of the all-or-nothing assignment method in the previous step that identifies origin/destination pair on the shortest routes where traffic counts occur and any origin/destination pair across the rows uses the traffic count its proportion is set to one or else zero that is called as "*Traffic count*" column for the particular link.

During the initialization phase, a starting initialization is done by assigning a minimum one trip for each OD pair. This first traffic count column is summed and is used to calculate initial trip volumes that can be called as "*Current Volume*" column (Wilson, 2012). The equation for each cell can be presented as follows (Eq. (3)):

$$\text{Current Volume} = \frac{\text{Traffic count value for the cell} * \text{Actual count of link a}}{\sum \text{Traffic count}} \qquad (3)$$

The summation of "*Current volume*" column will be equal to the Actual Count of link *a*. "*Updated OD trips*" column compares the current volume column with the initialized OD trips column and keeps the current volume when current volume is not zero, otherwise pull the value from the initialized OD trip value. This action can be carried out by using a set of if-else procedures (Wilson, 2012).

The next link *b* column is made up of proportion values for the origin/destination pairs that use the traffic count. In the column labeled "*b**" the proportion values in the column "*Traffic Count b*" are multiplied individually by the values in the "*Updated OD trips*" column. In the next column labeled "*Current Volume*" is a ratio of the observed/actual traffic count of link *b* compared to the sum of the "*b**" column that can be shown in the following Eq. (4) (Wilson, 2012):

$$\text{Current Volume} = \frac{\text{cell value of "b*"column} * \text{Actual count of link b}}{\sum \text{"b*"column}} \qquad (4)$$

The current volume column can be summed to yield a value which is the same as the observed traffic count. The Updated OD trips column can be calculated as stated above for link *a*. In similar fashion "*" column followed by "*Current Volume*" and "*Updated OD trips*" columns can be developed for the other links. "*Updated OD trips*" column is shaping up and refining across the link columns to produce a potential solution for the trip matrix (Wilson, 2012).

The iteration phase builds on what was done in the first initialization section. It starts off with the "*Traffic Count a*" column which is still populated with the same proportion values for the origin/destination pairs as before. The new column labeled "*a**" is calculated by multiplying each cell in the corresponding proportion value from the "*Traffic Count a*" column by the final "*Updated OD trips*" column found in the initialization spread sheet. The rest of the calculation will be the same as the initialization phase (Wilson, 2012).

If the number of iterations increase, the summation of "*" column for a specific link is approaching closer to the observed traffic count for that link and the difference between two final "*Updated OD trips*" values of consecutive iterations is getting lesser. In large network, it is more difficult to satisfy every observed count and it is wise to keep track of the difference of two consecutive iterated final Updated OD trip values. It can be noted that the count columns at the end of the spread sheet are likely to satisfy the observed counts. As this study is concern about the external stations, link columns must be rearranged in such a way so that all external counts will be placed at the end of the initialization and iteration spread sheets.

3.5. Formulation of Different Options

If standard initialization values show impractical results, means the final number of trips between zones are overestimated for certain origin destination pairs who are geographically located next to each other for external zones or are showing misleading amount of movements for an internal zone where negligible development inside the territory cannot justify the large movements. Therefore, based on the assessment of trip matrix developed under standard initialization, the following options can be formulated and can be reflected in the new initialization columns which can be called as constrained initialization columns:

- External OD pairs which are geographically closely located can be constrained by initializing their trip value as zero and will be unchangeable through iterations. It is reasonable to hypothesize that two external stations in close proximity to each other would be less likely to exchange through trips than two external stations at opposite sides of the urban area (Horowitz and Patel, 1999).
- Internal zones attracting or generating way more trips, that are causing huge deviation for other OD pair trips can be constrained by initializing all trips related to the Internal zone as zero and will be unchangeable till the end of the last iteration.
- Any combination of above two options can create other viable option.

Running initialization and iteration spread sheets for each option can be done and final trip matrices for each option can be stored for further analysis.

To avoid issues with the impractical numbers, essentially trips that enter the study and immediately turn around, the trips between external stations especially closely located and not passing through the study area (shown in Fig. 1), are constrained as zero trips. The following Table 2 details the clusters of external stations based on geographical setting.

Table 2
Geographical Clusters of External Stations

EE Origin/Destination (O-D) Clusters	O-D Movements
1, 2, 3, 4, 5	1-2, 2-1, 1-3, 3-1, 1-4, 4-1, 1-5, 5-1, 2-3, 3-2, 2-4, 4-2, 2-5, 5-2, 3-4, 4-3, 3-5, 5-3, 4-5, 5-4
1, 2, 12, 13	1-12, 12-1, 1-13, 13-1, 2-12, 12-2, 2-13, 13-2, 12-13, 13-12
5, 6, 7	5-6, 6-5, 5-7, 7-5, 6-7, 7-6
8, 9, 10, 11, 12, 13	8-9, 9-8, 8-10, 10-8, 8-11, 11-8, 8-12, 12-8, 8-13, 13-8, 9-10, 10-9, 9-11, 11-9, 9-12, 12-9, 9-13, 13-9, 10-11, 11-10, 10-12, 12-10, 10-13, 13-10, 11-12, 12-11, 11-13, 13-11

3.6. Evaluation/Validation of Options

The entire OD matrix contains E-I, I-E, E-E and I-I movements. As our study only concerned on the external movements, a matrix that only contains external movements was extracted and those trips were aggregated and used in the comparison purposes with the actual percent through trips and total trips of the external stations from the Bluetooth study. Each option can be examined after running certain number of iterations and following above procedures and the best option can be selected based on the minimum values of the total differences. To assess the trend of total differences by increasing the number of iterations, scatter plot can be drawn for the above two parameters. Parallel-constant lines will be found that represent no need of further iterations and thus the number of iterations can be settled at this steady state.

3.7. Results

The last iteration selected was number 151. The results are presented in the following Table 3.

Table 3

Results and Specific EE Movements

EE No	Through Trips from this study	Through Trips from Bluetooth	% Percent Through from this study	% Percent Through from Bluetooth	Difference in percent
1	115	91	11.01%	8.1%	-2.91%
2	85	21	11.01%	2.5%	-8.51%
3	272	65	14.97%	3.3%	-11.67%
4	631	3,506	2.56%	14.2%	11.64%
5	1	109	0.13%	11.1%	10.97%
6	885	863	6.93%	6.6%	-0.33%
7	273	216	3.15%	2.5%	-0.65%
8	8	14	1.79%	3.2%	1.41%
9	1,136	3,405	4.40%	13.2%	8.80%
10	40	361	0.59%	5.3%	4.71%
11	25	0	12.23%	0.0%	-12.23%
12	0	4	0.00%	2.3%	2.30%
13	21	538	0.28%	7.3%	7.02%

After looking at the Table 3, the two methodologies produce similar results for the majority of external stations. The external stations with the higher traffic volume, the traffic count methodology performs worse as it seems these trips are more likely distributed as E-I trips versus pass through. However, the difference in percent between the two studies is within +/- 13%. This discrepancy is similar to the values found when using regression models or synthetic procedures (Han, 2007; Anderson et al., 2006).

4. Discussion and Conclusion

Modlin's and Anderson's methods employ linear regression equations to estimate through trips. Discrete choice based model for smaller cities proposed by Martchouk and Fricker can ensure that the through trip percentages add up to100 percent

(Martchouk and Fricker, 2009). Recent research done by Talbot et al. (2011), developed a set of two logit models to estimate through trips for a wide range of study area sizes that requires a significant amount of data, including external survey data, traffic data, roadway data, demographic data, interaction score data, and measures of external station separation.

Unlike many previous models, this process does not require of obtaining sufficient amount of data for the study area. In previous models explanatory variables or predictor variables are must to form the final refined models. In this study only traffic counts and actual through trips are necessary database to find the best option with a cutoff point to run the number of iterations for estimating through trips, that is the uniqueness of this research. And traffic counts are accessible for any year that can easily be corresponded with the year of actual through trips. There is no doubt of its transferability for a wide range of study area sizes and this methodology can be applied in smaller or larger areas.

The case study demonstrated that the through trip values can be determined, within +/- 15 percent of actual values following the methodology. This research developed a step by step procedure to determine through trips pattern for any kind of study area by using minimal amount of existing data. The outcome of this research will be very useful for the urban areas where external survey is suspended or cannot be conducted due to lack of resources.

Acknowledgements

The authors acknowledge Alabama Department of Transportation (ALDOT) for their financial support. The authors would like to thank Assistant Prof. Jesmin Khan of Tuskegee University for her help in preparing the MATLAB script.

References

Abrahamsson, T. 1998. Estimation of Origin-Destination Matrices Using Traffic Counts - A Literature Survey, Laxenburg, Austria. 32 p.

Anderson, M.D.; Abdullah, Y.M.; Gholston, S.E.; Jones, S.L. 2006. Development of a methodology to predict through trip rates for small communities, *ASCE Journal of Urban Planning and Development*. DOI: http://dx.doi.org/10.1061/(ASCE)0733-9488(2006)132:2(112), 132(2): 112-114.

Bera, S.; Rao, K.V.K. 2011. Estimation of Origin-Destination Matrix from Traffic Counts: The Stat of the Art, *European Transport \ Transporti Europei*, 49(2011): 3-23.

Divison, T.D.-T.P.a.P. 2012. District Traffic Maps - 2011. Available from Internet: <http://ftp.dot.state.tx.us/pub/txdot-info/tpp/traffic_counts/2011/bry_base.pdf>.

Farnsworth, S.; Voigt, T.; Borchardt, D. 2011. *Alternative Methods for External Travel Surveys*. 20 p.

Han, Y. 2007. *Synthesized Through Trip Model For Small and Medium Urban Areas*, Raleigh, NC. 342 p.

Horowitz, A.J.; Patel, M.H. 1999. Through Trip Tables for Small Urban Areas: A Method for Quick Response Travel Forecasting, *Transportation Research Record: Journal of the Transportation Research Board*. DOI: http://dx.doi.org/10.3141/1685-08, 1685(1999): 57-64.

Martchouk, M.; Fricker, J.D. 2009. Through-trip Matrices Using Discrete Choice Models: A Planning Tool for Smaller Cities. Presented at the *Transportation Research Board 88th Annual Meeting*. 13 p.

Ortuzar, J.de D.; Willumsen, L.G. 1994. *Modelling Transport*, 2nd ed. New York, USA. 454 p.

Talbot, E.S.; Burris, M.W.; Farnsworth, S. 2011. Estimating Through Trip Travel without External Surveys. Paper Submitted for Publication and Presentation at the *Transportation Research Board Annual Meeting*. 104-111.

Wilson, A.G. 1970. *Entropy in Urban and Regional Modelling*, London. 178 p.

Wilson, J.P. 2012. *Evaluation of Origin/Destination Matrix Estimation Techniques to Support Aspects of Traffic Modeling*, Huntsville. 233 p.

Xie, C.; Kockelman, K.M.; Waller, S.T. 2010. A Maximum Entropy Method for Subnetwork Origin-Destination Trip Matrix Estimation, *Transportation Research Record: Journal of the Transportation Research Board*. DOI: http://dx.doi.org/10.3141/2196-12, 2196(2010): 111-119.