A TOUR-BASED URBAN FREIGHT DEMAND MODEL USING ENTROPY MAXIMIZATION

Qian Wang, *Assistant Professor, Department of Civil, Structural and Environmental Engineering*
University at Buffalo, the State University of New York, qw6@buffalo.edu
José Holguín-Veras, *Professor, Department of Civil and Environmental Eng., Rensselaer Polytechnic*
Institute, jhv@rpi.edu

1. Introduction

In recent years, one of the unique features of urban commercial vehicle movements, i.e., trip chaining behavior, has been receiving more attention. As it has been found (Ogden 1992; Holguín-Veras and Thorson 2003; Holguín-Veras and Patil 2005), commercial vehicles tend to make long tours composed of multiple trips that are interrelated according to the underlying logistic decisions. These long tours break down the typical assumption of the traditional four-step approach that assumes that trips are independent, and that the trips between an origin-destination (OD) pair are only related to the zonal attributes and the travel impedance of the corresponding OD pair. Therefore, new paradigms considering tours are needed for urban freight demand forecasting.

In order to address this feature, a handful of tour-based models have been formulated to model urban freight movements. Most of these approaches modeled commercial vehicle tours at the disaggregate level by solving a vehicle routing problem (Wisetjindawat et al. 2007; Donnelly 2007) or by using the probabilities generated by a set of discrete choice models (Stefan et al. 2005; Gliebe et al. 2007; Hunt and Stephan 2007). Wang and Holguín-Veras (2009) also proposed a hybrid micro-simulation modeling framework to generate goods-related vehicle tours that satisfy a known commodity flow origin-destination (OD) matrix in an urban freight network. These disaggregate models, on one hand, are capable to capture the underlying decision making process behind vehicle operations. On the other hand, they have some limitations that restrict their applications, particularly in large cities. These include the expensive procedures required for collecting travel diary data, the long computation time, and the strong reliance on the assumptions made about logistic operations. In contrast to the disaggregate approach, an aggregate tour-based approach can be used as an alternative way to model urban freight movements, considering its smaller data requirements, faster computational time, and less reliance on behavioral assumptions. However, very little has been made in this direction. The only related research found by the authors is from Maruyama and Harata (2005) who developed three types of combined network equilibrium models that take into consideration trip chaining behavior. However, since the modeling framework was proposed for overall traffic flows in the network, the issues and potential applications in forecasting freight demand were not discussed.

In this context, a powerful method, entropy maximization, is used to develop an aggregate tour-based freight demand model that is equivalent to the traditional gravity model. The main difference is that, while the gravity model estimates the flows from origin i to destination j; the new model does the same for an entire tour comprised of multiple stops. The entropy maximization assumes that in the absence of detailed information for individual commercial vehicle movements in an urban area, all the tour flows are assumed to be equally probable unless information is available to the contrary. Furthermore, of all the feasible ways to distribute commercial vehicle flows to the network, the most probable ones would be those that can be generated in the greatest number of ways under the constraints of the known aggregate information. Here, the known aggregate information includes the total number of commercial vehicle trips produced by (trip production) or attracted to (trip attraction) each node and the total impedance of the network. With the aggregate information as the input, the resulting entropy formulation is able to find the commercial vehicle tour flows that are consistent with the real-world freight movement patterns.

2. Problem Statement

Before the problem is stated, it is important to define the concepts of "tour," "tour flow," and "trip." A "tour" is defined as the sequence of nodes or locations visited by a vehicle, which starts and ends at the vehicle's home base. It is used interchangeably with "node sequence" in this paper. The number of vehicle journeys following a specific tour during a certain time period is referred to as "tour flow." An

individual vehicle movement connecting two consecutive stops is referred to as a "trip". These definitions are consistent with the previous publication (Holguín-Veras and Patil 2005).

This paper is aimed to estimate commercial vehicle tour flows in an urban area. Since tours are the physical node sequences where tour flows arise, they need to be specified before the tour flows can be estimated. To address this, a sequential modeling framework is proposed, which is composed of a tour choice model and a tour flow estimation model as shown in Figure 1.

The tour choice model is to produce a sufficient and effective set of tours that resemble trip chaining behavior of commercial vehicles. In this step, the challenges are: (1) how to generate the individual tours; and (2) since the number of possible tours is astronomical in a large network, how to select the tours in a way that the observed tour length distribution is represented while the enumeration of all alternative tours is avoided. For the first question, as have been discussed by Wang and Holguín-Veras (2008), discrete choice modeling can be used to generate tours by selecting stop locations and making the decisions about whether to return to the base or not for each tour. Regarding the second question, Wang (2008) developed a heuristic algorithm to screen tours based on the estimated behavioral choice models, and obtained a sufficient and effective set of tours for the Denver Metropolitan area which is used as the test case in this paper. Interested readers are referred to the two publications for more details.
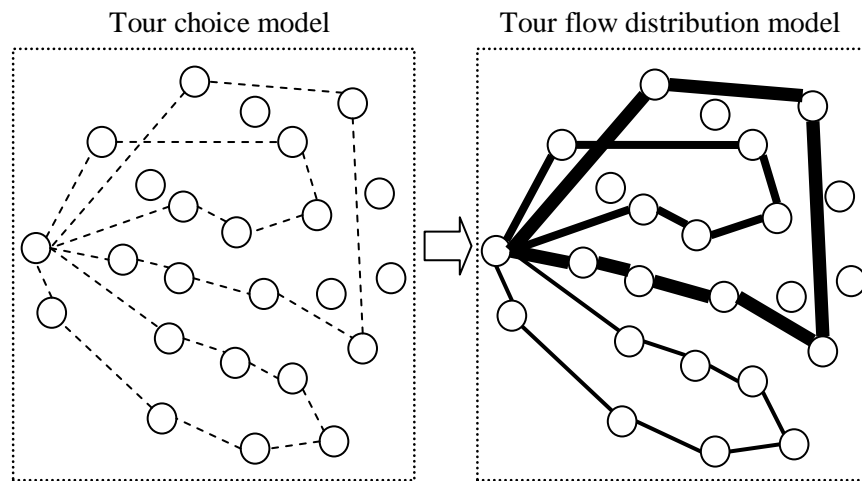

Figure 1. Tour-based urban freight demand modeling framework

After a tour set is estimated from the tour choice model, a tour flow distribution model is used to estimate the amount of commercial vehicle flows distributed to each tour, given the aggregate input information such as the number of trips produced by or attracted to each node and the total impedance in the area. In this stage, the major challenge comes from the enormous ways to distribute commercial vehicle flows. To address this issue, the entropy maximization is used to find the most likely set of tour flows that satisfy the input information. This will be the focus of this paper.

Several assumptions are made to develop the tour flow distribution model: (1) the nodes considered are the traffic analysis zones (TAZs) in urban areas; and (2) the tour set and the impedance of each tour are known from the tour choice model as the input.

3. Tour-based Entropy Maximization Model

The entropy maximization formulation developed here intends to solve the tour flow distribution problem of commercial vehicles in an urban area. For modeling purposes, three states are defined for an urban freight system of interest (see Table 1).

Table 1. States definitions

State	State Variable

Micro state	Individual commercial vehicle journey starting and ending at a home base (tour flow) by following tour m;
Meso state	t_m: The number of commercial vehicle journeys (tour flows) following tour m;
	O_i: Total number of trips produced by node i (trip production);
	D_j: Total number of trips attracted to node j (trip attraction);
Macro state	C: Total tour impedance in the commercial network;

Note: $m = \{1,2,...,M\}$, and M is the total number of tours.

The entropy formulation could be written as an equivalent minimization program as follows:

$$Min \qquad z = \sum_{m=1}^{M}(t_m \ln t_m - t_m) \tag{1}$$

Subject to:

$$\sum_{m=1}^{M} a_{im} t_m = O_i, \qquad \forall i \in \{1,2,...,N\} \qquad (\lambda_i) \tag{2}$$

$$\sum_{m=1}^{M} c_m t_m = C \qquad\qquad (\beta) \tag{3}$$

$$t_m \geq 0, \qquad \forall m \in \{1,2,...,M\} \tag{4}$$

Where:

λ_i = Lagrange multiplier associated with the ith trip production constraint;

β = Lagrange multiplier associated with the total impedance constraint.

The formulation above is composed of one objective function and four groups of constraints. Statement (1) indicates that the objective of this problem is to find the most likely ways to distribute tour flows. The first group of constraints, i.e., equation (2), is the trip production/attraction constraints. They indicate that the summation of the tour flows passing a node has to be equal to the total number of trips produced by or attracted to this node. Equation (3) is the total impedance constraint. It means that the summation of impedances of tour flows equals the total impedance in the network. The last group of constraints, i.e., equation (4), are the nonnegative constraints which mean the resulting tour flows should be equal to or greater than zero.

In order to understand the characteristics of the entropy maximization model, the first-order conditions (or Karush-Kuhn-Tucker conditions (KKT) conditions) and the second-order conditions (Hessian matrix), are obtained for the entropy model. The first-order conditions result in the optimal tour flow distribution model as below:

$$t_m^* = \exp(\sum_{i=1}^{N} \lambda_i^* a_{im} + \beta^* c_m) \tag{5}$$

It indicates that the number of tour flows loaded on a node sequence (t_m^*) is an exponential function of the Lagrange multipliers associated with the trip productions/attractions of nodes along that tour and the tour impedance (c_m). This indicates that the tour flows are both affected by the level of trip generation and the tour cost, and the impacts are quantified by the corresponding Lagrange multipliers (λ_i and β). Meanwhile, since the Hessian matrix of the objective function is positive definite, the objectives function is convex. Since the constraints are linear, the overall entropy maximization formulation is convex, which indicates the uniqueness of the optimal tour flow solutions.

4. Case Study

In order to test the practicality and validity of the tour-based entropy model, it was applied to the commercial vehicle tour data collected in the Denver metropolitan area between 1998 and 1999. 613 complete tours and the corresponding travel times were observed in the data, representing a total of 65,385 journeys made on these tours per day. Due to the complexity of the entropy model, an algorithm developed for solving large-scale entropy maximization problems, i.e., the primal-dual interior method for optimization programs with convex objectives (PDCO) (SOL 2005; Paige and Saunders 1975; Saunder 1996; Saunder and Tomlin 1996) was used to solve the problem. It took less than one minute to calculate the optimal tour flows.

The estimated tour results clearly show the approach's accuracy and efficiency of the approach. As shown in Figure 2 and
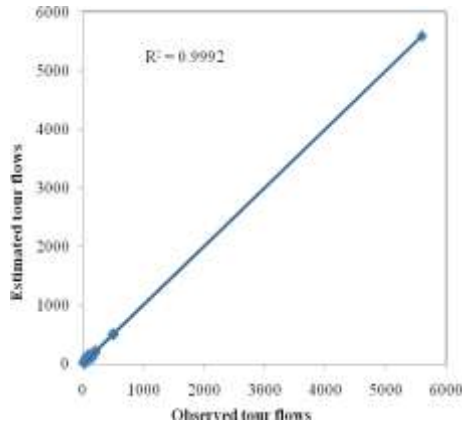

Figure 2: Estimated vs. observed tour flows

Table **2**, the estimated tour flows closely match the observed values, and the mean absolute percentage error was only 6.71%. Equally important is that the parameters of the model, and the resulting trip length distributions were conceptually valid. As revealed by the negative sign of tour travel time (-0.000228), longer the tour length is, less likely tour flows will be distributed to that tour, which is consistent with the fact that commercial vehicles prefer short tours. These results provide a strong indication of the model's potential as the underlying travel demand forecasting model for commercial vehicle movements in urban areas.

Table 2. The MAPE and the estimated impedance-related Lagrange multipliers

Estimated Results	Entropy model
MAPE	6.71%
Tour-time-related Lagrange multiplier (β)	-0.000228

5. Conclusions

This paper discusses a tour-based entropy maximization model that is intended to predict the commercial vehicle tour flows given the aggregate demand information such as the number of trips produced by or attracted to each node, and the impedance to travel in urban areas. The first and second order conditions were derived to gain insight from the entropy maximization formulation. The first-order conditions derive an optimal tour flow distribution model that shares the similar functional form as the traditional gravity model but focuses on tours instead of trips. This tour flow distribution model shows that the flow of commercial vehicles traveling along a given tour is a function of the Lagrange multiplier associated with the number of trips produced by or attracted to each node along that tour, and the tour impedance. The second-order conditions indicate the uniqueness of the optimal tour flow solutions. A case study in the Denver metropolitan area shows the accuracy and efficiency of this approach: the estimated tour flows closely match the observations with the mean absolute percentage error as 6.71%. In addition, as indicated by the tour flow distribution model derived, longer the tour length is less tour flows will be distributed to that tour.

6. Acknowledgements

7. References

Donnelly, R. (2007). A Hybrid Microsimulation Model of Freight Flows. *Proceedings of the 4th International Conference on City Logistics*, July 11-13, 2007. Taniguchi, E. and R.G. Thompson (Ed.). Crete, Greece, Institute for City Logistics, pp: 235-246.

Holguín-Veras, J., and Thorson, E. (2003). Modeling Commercial Vehicle Empty Trips with a First Order Trip Chain Model, *Transportation Research Part B*, Vol. 37, pp. 129-148.

Holguín-Veras, J., and Patil, G.R. (2005). Observed Trip Chain Behavior of Commercial Vehicles, *Transportation Research Record: Journal of the Transportation Research Board*, No. 1906, pp. 74-80.

Hunt, J.D., and Stefan, K.J. (2007). Tour-based Microsimulation of Urban Commercial Movements. *Transportation Research Part B*, Vol. 41, Issue 9, pp. 981-1013.

Gliebe, J., Cohen, O., and Hunt, J.D. (2007). A Dynamic Choice Model of Urban Commercial Vehicle and Person Activity Patterns. *The 86th Transportation Research Board Annual Meeting*, CD-ROM.

Maruyama, T., and Harata, N. (2005). Incorporating Trip-Chaining Behavior into Network Equilibrium Analysis. *Transportation Research Record*, No. 1921, pp.11-18.

Ogden, K.W. (1992). *Urban Goods Movement: A Guide to Policy and Planning*. Ashgate, Hants,UK.

Paige, C., and Saunders, M.A. (1975). Solution of Sparse Indefinite Systems of Linear Equations. *Siam J Numer Anal*, Vol. 12, No. 4, pp. 617- 629.

Saunders, M.A. (1996). Cholesky-based Methods for Sparse Least Squares: The Benefits of Regularization. In L. Adams and J. L. Nazareth (eds.), *Linear and Nonlinear Conjugate Gradient-Related Methods*, SIAM, Philadelphia, pp. 92-100.

Saunders, M.A., and Tomlin, J.A. (1996). Solving regularized linear programs using barrier methods and KKT systems, Report SOL 96-4, Dept of EESOR, Stanford University.

Stefan, K.J., McMillan, J.D.P., and Hunt, J.D. (2005). Urban Commercial Vehicle Movement Model for Calgary, Alberta, Canada. *Transportation Research Record: Journal of the Transportation research Board*, No. 1921, pp. 1-10.

SOL (Systems Optimization Laboratory) (2005). Free Software Download. Accessible at: http://www.stanford.edu/group/SOL/software/pdco.html. Accessed: July 10, 2008.

Tomlin, J.A. (1971). A Mathematical Programming Model for the Combined Distribution-Assignment of Traffic. *Transportation Science*, Vol. 5, pp. 122-140.

Wang, Q., and Holguín-Veras, J. (2008). Investigation of Attributes Determining Trip Chaining Behavior in Hybrid Microsimulation Urban Freight Models. *Transportation Research Record: Journal of the Transportation research Board*, No. 2066, pp. 1-8.

Wang, Q. (2008). Tour-based Urban Freight Travel Demand Models. Ph.D. dissertation. Department of Civil and Environmental Engineering, Troy, NY, Rensselaer Polytechnic Institute.

Wisetjindawat, W., Sano, K., Matsumoto, S., and Raothanachonkun, P. (2007). Micro-Simulation Model for Modeling Freight Agents Interactions in Urban Freight Movement. *The 86th Transportation Research Board Annual Meeting*, CD-ROM.