CE298 – Project 2 (Literature Review)

Tour Based Freight Modeling

(Focused on Entropy Maximization Aggregate Models)

- *Anupam Srivastava*

Introduction

The movement of commercial vehicles constitutes a sizeable component of total vehicle movement on urban road networks. Freight demand modeling, as a result, has become a topic of increasing interest in recent years. Unlike modeling of passenger trips, which has been studies in great detail over the years, the modeling of commercial vehicles is still a new topic with many unanswered questions. The higher complexity in the behavioral nature of the problem demands greater insight into the modeling technique used for freight demand modeling, where traditional 4-step models might not be sufficient.

There have been several studies recently exploring the trip chaining behavior of vehicles. Studies have shown that commercial vehicles tend to make long tours composed of multiple connected trips that are interrelated based on various logistics choices. A traditional 4-step model that estimates independent trips between pairs of Origins and Destinations based on corresponding zonal attributes and impedances alone is not able to capture this chaining behavior well. Tour-based models, that use tours instead of trips as the unit of movement being modeled, have thus been developed to address this concern. Vehicle tours are either modeled at the individual vehicle level through disaggregate tour models, or through aggregate models estimating total flow along possible tours(6,12,9). At the disaggregate level, the models are either based on solving a variation of the vehicle routing problem (1,11), or use a probabilistic discrete choice model (3,5,7).

The tour-based freight models as a class of freight modeling techniques are somewhat similar to logistics models that model the movement of commodities, and often grouped together with activity based models. The distinction between tour-based and activity based models is in fact very hazy. Unlike their logistics model counterpart, they do not directly model the movement of commodities, and thus are not useful in assessing impacts of policies affecting commodity market behavior and flows. However since they generate actual truck volumes for various routes, they are directly equipped (and far better so than any other modeling technique) to answer vehicle based policy questions such as freight HOV and other road infrastructure related policies etc. This makes them a very useful tool for freight modeling.

Tour-based Models

While 4-step demand models use vehicle trips as the fundamental unit of analysis, tour-based models, as their name suggests, instead model vehicle tours. In doing so, they are able to capture the interrelated chaining of trips forming a single tour. The distinction between trips and tours is an important aspect in understanding the strengths of tour-based models. While a vehicle trips is any movement of a vehicle between two nodes (origin to destination), a vehicle tour is a representation of a sequence of interrelated movements made by a vehicle. Usually a tour starts at a home node, visits an ordered sequence of nodes, and terminates returning to the home node (although not required to). A typical representation of a tour along with the corresponding breakup of component trips is shown in Figure 1.

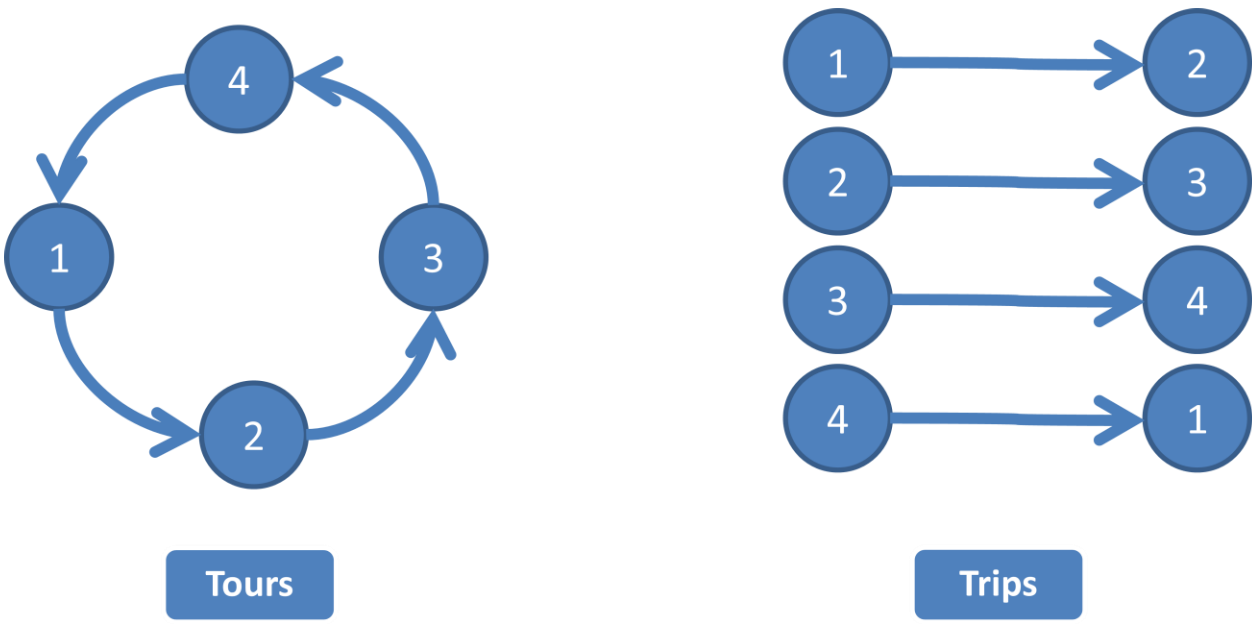

Figure 1: Defitions for tours and trips

Tour-based modeling techniques are usually very efficient at capturing the vehicle flow attributes over a network, capturing the constituent vehicle movement patterns. This makes them much more useful at analyzing affects that vehicular policies might have on the overall freight system for a region. Further, they are able to provide a higher detail of information as compared to some other more aggregate estimation models. Most tour-based models that attempt at modeling an individual commercial vehicle's movements, however, can be extremely data intensive requiring access to very detailed 'travel-diary' like data. The intensive data usage, also add complexity to the models, also making them time intensive. Further, discrete choice based models usually assume certain logistics behavior choice paradigms. It is not trivial to prove that such underlying assumptions in the model are always justified. Further, while being extremely useful in modeling under scenarios that explore the effects vehicular policy changes might have, tour-based models can only to a moderate degree, predict responses to commodity based policy changes. Such modeling is done through associating individual vehicles within the model to classes of commodities that they might be most likely to carry, in order to estimate the corresponding commodity flows. Tour-based models, though with some unarguable strengths, are most often not practically applicable with the current access to data collection resources and modeling tools. Only two of the models discussed have been actually implemented for a real network, Hunt and Stefan's model being used for Calgary, Alberta, Canada, and Gliebe's model that was implemented for Ohio. However, the evolving knowledge and interest in the field, and continuing efforts towards gathering better freight data, show promise for tour-based models since they are indeed able to capture certain behavior aspects of freight transportation that the more traditional models are not as equipped to handle. We see next how a certain class of tour-based models can handle some of the concerns discussed above.

Aggregate and Disaggregate Tour-Based Models

Tour-based models can be broadly classified under disaggregate, and aggregate tour models. The disaggregate models attempt to model the movement of individual vehicles in a network based on either vehicle routing techniques, or through discrete choice models. Such models are often presented in a micro-simulation based environment to show the individual vehicle movements. Aggregate characteristics such as link and route flows are calculated by summing up the appropriate individual vehicle movements. Aggregate tour-based models, as the name suggests, sacrifice the detail of vehicle movements at individual levels, and instead model total route flows over links in a network. In using an aggregate approach, such models overcome the complexity and data intensiveness issues that are characteristic of disaggregate tour models. Further the inherently required modeling assumptions are shifted from being behavior choice assumptions of logistics managers and drivers, to network based aggregate route / tour choice assumptions which are expected to have a higher precision. Table 1 shows a summary of the distinctions, advantages and disadvantages of aggregate and disaggregate tour models against each other.

Table 1: Aggregate and Disaggregate Models

	Aggregate	Disaggregate
Differences	• Tour distribution • Less Sensitive to time window	• Individual tour with time window
Pros	• Smaller data requirements • Faster computation time	• Capable of capturing the underlying decision making process behind vehicle operations

Cons	• Less reliance on behavioral assumptions	• Expensive procedures such as collecting travel diary data • Long computation times

Source: Soyoung You, PhD Dissertation, 2012

Disaggregate Tour-Based Models

Disaggregate tour-based models look at trying to model the movement patterns of each individual commercial vehicle being modeled individually. Such models are usually based either on solving a variation of the vehicle routing problem (such as Wisetjindawat et. al. 2007 [11], Donnelly 2007 [1], and Figliozzi 2007 [2]), or using probabilistic discrete choice models (such as Hunt et. al. 2005 [7], and Gliebe et. al. 2007 [3]). While disaggregate tour models provide a very high degree of detail about vehicle movements and patterns, they often have a heavy data requirement. Even so, Hunt's model has been implemented in the city of Calgary and has been discussed often in other works.

Gliebe et. al. 2007

Gliebe et. al. in 2007 (3) proposed a dynamic activity choice model as a disaggregate tour-based freight modeling scheme. The model is run in a disaggregate micro-simulation environment. The model uses A dynamic choice schemes to incrementally build tours accounting for activity delays, elapsed time, time of day etc. in deciding the next activity and location selection. The overall scheme, called the Disaggregate Commercial Model (DCM) was applied to the Ohio statewide model system. DCM does not model fixed route, or patrol-type movements such as taxis, garbage trucks, school busses, mail trucks, and police cars.

Figure 2 shows the overall structure of the DCM. As seen through the figure, the DCM first generates worker travelers, then assigns vehicles to the traveler, followed by assigning a starting time before going into the Dynamic Activity Pattern Generation. The input to the system is the zonal land use data and the survey data for calibration of the discrete choice models implemented. The first 3 stages of traveler generation, vehicle assignment, and starting time assignment are done using binary choice formulation, multinomial logit model, and probabilistic drawing from an empirical distribution respectively. At the end of the dynamic activity pattern generation model, DCM produces the individual vehicle trips at the disaggregate level.

The dynamic activity choice model, incrementally assigns activities and activity locations to the vehicle-traveler. The choices are made using a discrete choice modeling scheme based on probabilities calculated using various factors such as time of the day, and activity delays. Once the initial assignments are done, each subsequent assignment is a choice between the various available alternative activities, as well as a "stay" activity which retains the vehicle at the current activity and location. Thus, an individual switches from one activity to another only when an alternative becomes more attractive, as opposed to older models that force fixed activity delay times at each activity location and force an alternative choice to made at the end of the duration. In the current model, the delay at each location can be modeled intrinsically based on when its value diminishes.

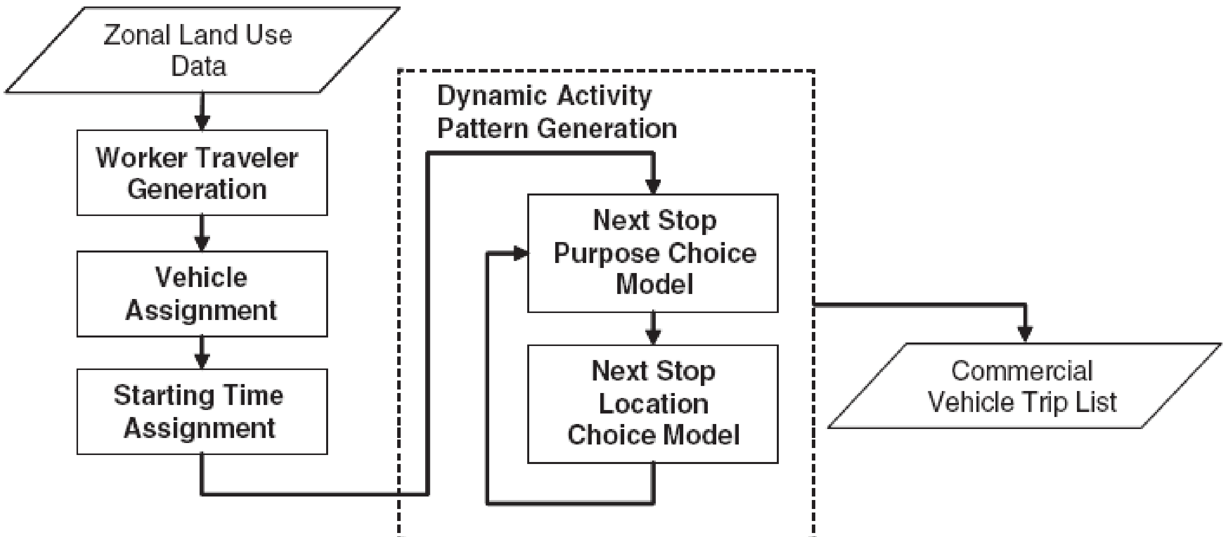

Figure 2: Structure of DCM
Source: Gliebe et. al. 2007 (3)

The DCM model was applied to a case study for Ohio especially for the 'next stop purpose' estimation aspect of the model. The following list provides a summary of the results from the statistical tests presented by the authors:

- Log likelihood at 0: -16,149
- Log likelihood constants only: -8,754
- Final log likelihood: -7,270
- ρ^2 (relative to 0 coefficients): 0.550
- ρ^2 (relative to constants): 0.170
- ρ^2 ('transition of state' constants alone): 0.380

Stefan and Hunt 2005

Stefan, McMillan, and Hunt in 2005 (7) developed a disaggregate tour-based model involving micro-simulation of commercial vehicle traffic and implemented the model for the city of Calgary, Alberta, Canada. Their work was the predecessor and the foundation work on which Gliebe's model (discussed above) was designed. Hunt's model has a very similar structural design as the model later proposed by Gliebe. The flow structure for the model is shown in Figure 3. The model was specifically designed keeping urban commercial freight movement in mind and was appropriately named CVM for Commercial Vehicle Movement model.

The first step of the model is to generate tours, followed by assigning vehicles and tour purposes, and then by assigning the tour start times for each vehicle tour. The model then does an iterative run of the 'Stop purpose' generation and 'Stop location' selection phases to incrementally build up the entire tour for the vehicle, adding one activity node at a time. Unlike Gliebe's DCM model, the CVM model selects a fixed stop duration associated with a chosen activity from a distribution of stop times for the given activity, and thus proceeds to selection of a new 'stop purpose' once the stop duration clock has run out.

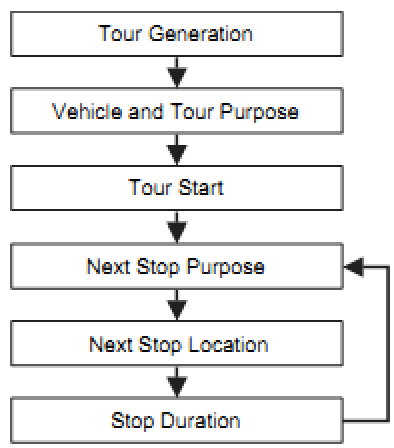

Figure 3: Process flow for Calgary CVM model

The CVM model was implemented for the city of Calgary, Alberta, Canada with success. Being among the earliest freight tour based models, and its successful implementation in a real world network has made the CVM model among the most widely studied tour-based models for freight.

Other disaggregate models

Figliozzi 2007

Figliozzi in 2007 (2) proposed a model for freight based on solving the vehicle routing problem to analyze impacts from congestion or technological changes to the freight system. One of his findings, interestingly, was that the percentage of empty vehicle trips did not influence the overall efficiency of the generated tours for the freight network.

Donnelly 2008

Donnelly in 2008 (1) proposed a commercial vehicle tour model that is developed from solving the traveling salesman problem of the empty backhaul trips of vehicles.

Aggregate Tour-Based Models

An aggregate approach to tour-based models offers advantages over the most concerning weaknesses of a disaggregate tour model: the data and time complexities. An aggregate approach to the modeling chooses to model the aggregate tour flows for candidate tours (sequence of nodes) instead of trying to model each individual vehicle's movement. In doing so, the data requirements are reduced by a great extent as is the complexity of the model itself. There have only been limited efforts towards this approach however. The approach was first used by Maruyama and Harata in 2005 (6) who were motivated by trip chaining behavior in traffic. This study developed three types of combined network equilibrium models that accounted for trip chaining. The study however, was aimed at modeling general traffic behavior on a network (passenger as well as freight). The study, therefore, did not discuss the issues and potential applications of the model specifically for freight demand forecasting / modeling. Holguin-Veras, as we see next, proposed the freight specific aggregate tour based modeling technique through their entropy maximization model formulation.

Entropy Maximization Models

Holguin-Veras in 2009 (9) introduced an aggregate level tour-based freight demand modeling technique based on an entropy maximization model. Entropy maximization models (minimum information theory), introduced by Wilson in 1970 (10), has often been used as a trip matrix estimation method for transportation modeling. Holguin-Veras used the same definitions for states at the micro, meso, and macro levels, as Wilson did in his original study (also shown in Figure 4): the micro state represents each individual vehicle trip that constitutes the tour flow for a given tour m, the meso state defines the number of trips between each OD pair, and the macro state modeling the total number of trips produced by or attracted to each zone in the network.

In the entropy maximization scheme, all micro states are assumed to be equally probable under lack of knowledge otherwise. The meso states are however not all equally probable, and the probability is related to the number of possible ways of generating a meso state under the given macro state constraints of total productions and attractions at the zonal level. The entropy is defined as the number of ways of generating the meso states in the network.

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;
Macro state	O_i : Total number of trips produced by node i (trip production); D_j : Total number of trips attracted to node j (trip attraction); Formulation 1: C : Total tour impedance in the commercial network; Formulation 2: C_T : Total tour travel impedance in the commercial network;

Figure 4: Definition of states

Based on the above definition of the entropy, the entropy can be written as:

$$Max\ \ W = C_T^{t_i} * C_{(T-t_2)}^{t_2} * \dots = \frac{T!}{\prod_{m=1}^{M} t_m!}$$

With the following constraints for total origin (O_i) and destination (D_j) counts at macro level, and travel impedance (C):

$$\sum_{m=1}^{M} a_{im} t_m = O_i, \qquad \forall\, i\ \in\ \{1, 2, \dots, N\}$$

$$\sum_{m=1}^{M} a_{jm} t_m = D_j, \qquad \forall\, j\ \in\ \{1, 2, \dots, N\}$$

$$\sum_{m=1}^{M} c_m t_m = C$$

$$t_m\ \geq 0, \qquad\qquad \forall\, m\ \in\ \{1, 2, \dots, M\}$$

The objective function is then simplified by first taking the natural logarithm of the formulation to get:

$$Max\ \ z' = \ln(W) = ln(T!) - \sum_{m=1}^{M} \ln(t_m!)$$

And then identifying and dropping the constant term ln(T!), and using Stirling's formulation to expand the logx factorial term into the more workable equivalent (log x! = xlnx – x). The maximization of the negative of the remaining term is then converted to the minimization of the value of the term without the signage. The new objective function becomes:

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

In the formulation, the objective function reflects the maximization of the entropy and thus finding the most likely ways to distribute tours under the constraints. The first and second sets of constraints require that the summation of all tour flows passing a node have to match the total number of trips produced (O_i) and attracted (D_j) to the node. The third set of constraints show that the summation of impedances for the tour flows give the total tour impedance (C) for the network. The final constraints the non-negativity constraints on the resulting tour flows.

Holguin-Veras pointed out, for definitions of tours where the first and the final nodes visited are both the home node (all tours end where they started), either the constraint for the total origins, or the constraint for the total destinations at a given zone can be dropped since they are redundant. Holguin-Veras further proposed two variations of the entropy maximization formulation for aggregate modeling of freight tours based on the travel impedance definition being used.

Tour Based Formulation 1

If only the total tour impedance is considered (thus accounting for all time spent during the travel and during the activity itself), we get the following formulation:

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

$subject\ to$:

$$\sum_{m=1}^{M} a_{im} t_m = O_i, \qquad \forall\, i \in \{1, 2, \dots, N\}$$

$$\sum_{m=1}^{M} c_m t_m = C$$

$$t_m \geq 0, \qquad \forall\, m \in \{1, 2, \dots, M\}$$

Tour Based Formulation 2

The alternate formulation uses the distinction between travel impedances (C_T) and handling impedances (C_H) through using two impedance constraint sets. Thus, using total tour travel impedance and total tour handling impedances separately, we get the following formulation:

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

$Subject\ to$:

$$\sum_{m=1}^{M} a_{im} t_m = O_i, \qquad \forall\, i \in \{1, 2, \dots, N\}$$
$$\sum_{m=1}^{M} c_m t_m = C_T$$
$$\sum_{m=1}^{M} c_m t_m = C_H$$
$$t_m \geq 0, \qquad\qquad \forall\, m \in \{1, 2, \dots, M\}$$

Holguin-Veras applied their two formulations for entropy maximization based freight tour models to a case study done for the Denver metropolitan area. Data required for the study was available through commercial vehicle travel diary survey records (done as part of a Travel Behavior Inventory survey conducted in 1998-1999). The vehicle travel diaries were collected for 832 vehicles of which 502 made at least one trip on the survey day. The calibration of the models presented were shows to take only a few seconds thus suggesting that the time complexity of such an aggregate model is far smaller than that for a disaggregate tour-based model. The mean absolute percentage errors of estimated tour flows recorded for the case study were 6.71% for formulation 1 and 6.61% for formulation 2. A summary of the results is shown in Table 2.

Table 2: MAPE and calibrated Lagrange multipliers

Estimated results	Formulation 1	Formulation 2
MAPE	6.71%	6.61%
Cost-related Lagrange multipliers:		
Tour-time-related Lagrange multiplier ((β)	-0.000228	/
Tour-travel-time-related Lagrange multiplier (β_1)	/	-0.00190
Tour-handling-time-related Lagrange multiplier (β_2)	/	0.000183

<u>Soyoung You</u>

You in 2012, (15) developed a modification to the entropy maximization formulation for freight tour models. Observing the behavior of drayage trucks for the twin ports of Los Angeles and Long Beach in California, she observed that drayage trucks tend to have more than one tour per day. Further, many tours were observed to contain repetitive patterns, thus generating numerous similar (but not perfectly identical) tours. Further, she recognized the limitation of the formulation approach used by Holguin-Veras that fails to capture the differences due to the exact ordering of nodes visited within a tour. While the Holguin-Veras formulation captures the combination of nodes visited in a tour, the order of visit is not distinctly captured. If instead, a path based approach was used as opposed to the node-based approach, the correct sequencing of the tour nodes would also be captured.

Using the same definitions for tours, and for the micro, meso and macro states, as before, we get the following formulation, with an added constraint:

$$Max\ W = \frac{x!}{\prod_{j=1}^{J} x_j!}, \qquad x_j \geq 0, \qquad \forall\, j\, \in \{1, 2, 3, \dots, J\}$$

Simplified to give the following objective function as before:

$$Min\ z = \sum_{j=1}^{J} \left(x_j \ln x_j - x_j\right), \qquad x_j \geq 0, \qquad \forall\, j\, \in \{1, 2, 3, \dots, J\}$$

$$Subject\ to,$$

$$\sum_{j=1}^{J} a_{ij} x_j = O_i\,, \qquad \forall\, i\, \in \{1, 2, 3, \dots, N\}$$

$$\sum_{j=1}^{J} b_{ij} x_j = D_i\,, \qquad \forall\, i\, \in \{1, 2, 3, \dots, N\}$$

$$\sum_{j=1}^{J} l_{kj} x_j = L_k\,, \qquad \forall\, k\, \in \{1, 2, 3, \dots, K\}$$

$$\sum_{j=1}^{J} c_{Tj} x_j = C_T$$

$$\sum_{j=1}^{J} c_{Hj} x_j = C_H$$

While the first and the last two set of constraints are the same as those proposed earlier, the third constraint is new. Here, l_{kj} is the number of times that an OD pair k is included in a given tour j, and L_k is the total number of truck trips between a given OD pair k. Since the equivalent OD trip breakups are also matched to the corresponding total number of trips between an OD pair, the ordering of the nodes within the tours is also accounted for in the constraints. Thus, the distinction between the tour A-B-C-D and the tour A-C-B-D is maintained. This is further illustrated in Figure 5.

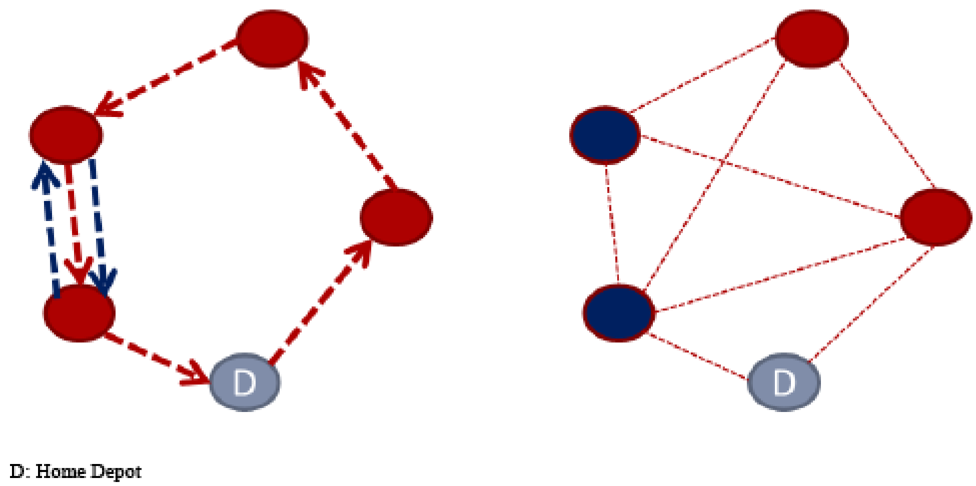

Figure 5: Path based vs Node based formulation

Discussions

There have been a few attempts at developing tour-based freight models at both the disaggregate and the aggregate levels. Tour based models suffer from being data intensive due to their higher level of resolution in the results produced, and the corresponding higher complexity, but are among the only models that can actually model movement of vehicles and flows of vehicles over a network. Thus, they might be the only tool available for studying impacts from certain policies such as changing HOV restrictions for heavy vehicles or restricting access to certain roads based on time of day etc.

Disaggregate tour based models offer a very detailed insight into the movement of individual vehicles. Such a model can be easily integrated into a micro-simulation software to study the volumes and congestion effects of these freight vehicles on links and on networks. The detailed model however comes at the cost of being highly data and time intensive in nature. Lack of availability of freight data, as is the case most of the times, would usually severely restrict real world applications of such disaggregate tour based models. Aggregate tour based models are over to overcome this hurdle to some extent, by reducing the data requirement by switching to a focus on tour flow volumes instead of movement of individual vehicles. Such a model, is still able to model the flow behaviors of freight movement (albeit at a higher level of aggregation) and thus can still be as useful in studying policy impacts.

It is important to understand the underlying concepts behind these models as well however, to see where they might be lacking, or where they might be making inherent assumptions that might not always hold. Discrete choice models that model tours at a disaggregate level, usually imply that a choice is being made at each incremental step for deciding the next activity and destination. However, such a discrete choice model is only applicable when applied to an individual who makes a choice not influenced directly by the choices of others (except indirectly due to the effect that others' choices might have on the impedances observed by the user). It might be questionable whether this is truly the case at hand in freight movements. One would assume, that an individual (driver) usually does not have unchallenged say in choosing his/her tour, and is instead guided greatly by a logistics / dispatch manager. Thus the choice is made by the carrier's operations unit instead of individual vehicles. If the operation unit was optimizing each vehicles' tours independent of other vehicles at the carrier's disposal, then the behavioral assumption might still hold. In such a case, the decision making process would simply be transferred from the vehicle driver to the operator. However, one would also expect that the operator is optimizing the tours considering all vehicles at disposal in his mind, and thus the discrete choices are no longer independent of one another. This might seriously jeopardize the assumptions for such an implementation.

It might also be important to note that production / attraction models that are used to derive OD matrices are based on modeling the production and/or attraction based on socio-economic factors characteristic to a zone. When looking at vehicle tours instead of trips however, the notion of production / attraction nodes, versus origin / destination nodes might become complex and it might not always be possible to identify the appropriate production and attraction for each trip comprising the tour. Say a tour involves going in the following node sequence: a-b-c-d-c-d-a. Which nodes should be treated as the production / attraction nodes might depend on how the movements themselves happen and where the loading / unloading of the goods are happening, a pattern that need not be unique to the tour sequence. Further, in the models proposed, there is a possibility of losing the directionality information for a trip as a component of the tour.

As with any other modeling technique, care should be taken to ensure that the model being used for a certain application, and the principles and assumptions it is built on, is valid for the purposes that are be realized by the model.

References

1. Donnelly, R., 2008. *A hybrid micro-simulation model for freight flows*. In: Innovations in City Logistics, Nova Science Publishers Inc., pp. 249-259.
2. Figliozzi, M.A., 2007. *Analysis of the efficiency of urban commercial tour: Data collection, methodology, and policy implications*. Transportation Research Part B, 41(9), pp. 1014-1032.
3. Gliebe, J., Cohen, O., Hunt, J.D., 2007. *Dynamic choice model for urban commercial activity patterns of vehicles and people*. Transportation Research Record: Journal of the Transportation Research Board, 2003, pp. 17-26.
4. Greaves, S., Figliozzi, M.A., 2008. *Commercial vehicle tour data collection using passive GPS technology: Issues and potential applications*. Transportation Research Record: Journal of the Transportation Research Board, 2049, pp. 158-166.
5. Hunt, J.D., Stefan, K.J., 2007. *Tour based microsimulation of urban commercial movements*. Transportation Research Part B: Methodological, 41(9), pp. 981-1013.
6. Maruyama, T., Harata, N., 2005. *Incorporating trip-chaining behavior into network equilibrium analysis*. Transportation Research Record: Journal of the Transportation Research Board, 1921, pp. 11-18.
7. Stefan, K.J., McMillan, J.D.P., Hunt, J.D., 2005. *Urban commercial vehicle movement model for Calgary, Alberta, Canada*. Transportation Research Record: Journal of the Transportation Research Board, 1921, pp. 1-10.
8. Wang, Q., Holguin-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, 2066, pp. 1-8.
9. Wang, Q., Holguin-Veras, J., 2009. *Tour based entropy maximization formulations of urban freight demand.* In: TRB 2009 Annual Meeting CD-ROM, 99th Transportation Research Board.
10. Wilson, A.G., 1970. *Entropy in urban and regional modeling*. London: Pion.
11. Wisetjindawat, W., Sano, K., Matsumoto, S., Raothanachonkun, P., 2007. *Micro-simulation model for modeling freight agent interactions in urban freight movement*. In: CD Proceedings, 86th Annual Meeting of the Transportation Research Board.
12. You, S., 2012. *Methodology for tour-based truck demand modeling*. PhD. Dissertation Report