A Stochastic Multi-modal Freight Transport Assignment Model with Random Coefficients

Otto Anker Nielsen

Centre for Traffic and Transport, Technical University of Denmark
Building 115, st. tv, Bygningstorvet
2800 Kgs. Lyngby, Denmark
oan@ctt.dtu.dk

1 Introduction

Freight transport assignment is often done by purely deterministic, user or system equilibrium solution methods. The models may include a multi-class structure, i.e. allowing different weights on the attributes in the cost function for different modes and freight classes. However, few models follow a general utility maximum framework that also allow for description of overlapping routes as well as heterogeneities within freight groups, i.e. random coefficients in the utility function. The paper present a stochastic schedule-based freight transport assignment model that allow for both an error term dealing with the overlapping route problem, random coefficients dealing with heterogeneities, and cost functions dealing with capacity problems. This is done within different choice levels, each consisting of a set of possible mode and route alternatives.

The paper presents the results and recommended modelling framework resulting from a Danish pre-study for a national model (Nielsen et.al. 2003a). The recommended approach from this study was to consider freight transport assignment, mode chains and mode choice as a joint problem. This distinguishes between decisions made by transport buyers, choices of multi-modal chains, and finally path choice through a network with a given set of a priori modes.

The model focuses on the Zealand and Skåne regions in East Denmark respectively Southern Sweden. The area includes the capital region of greater Copenhagen in Denmark, and the Malmö region of Skåne with about 3 millions inhabitants together. The two regions are separated by sea with two competing ferry crossings and one toll bridge/tunnel. The later includes a rail link. The road network has several alternative routes within each region, while the rail network has only a few alternative routes – in many cases just one.

2 Choice levels

The Danish Prestudy recommended a 3-level hierarchy of mode and route choice models to describe:

1. Decisions at the level of transport buyers, e.g. door to door truck transport versus multimodal transport. Transport buyers' decisions can in most cases be dealt with by discrete choice models, e.g. a nested logit model. Some choice sets are almost given already at this level, e.g. coal import to power plans by ships, while others are more open, e.g. multimodal truck-rail transport versus direct trucks.

2. Choices in multi-modal networks between terminals. At this level only connections (route segments) between terminals are described (Nielsen & Frederiksen, 2001b & 2001c). Decisions at this level are typically taken by the transport firm, and can be discrete choices, pure route choices, or a combination of the two.

3. Choices of paths between terminals. This may be described by a detailed route choice model. For trucks, decisions are typically taken by the driver (except cases with restricted/forced routes), while trains are typically directed by the rail administration.

3 Choice set generation

Predefined choice sets has e.g. been implemented in the Danish Storebælt, Øresund and Femer Belt forecast models which combine discrete choice in specific corridors in conjunction with all-or-nothing in between. In all these cases, the choice set was pretty obvious and of a reasonable size. Such an approach is adopted in the present model, where the upper level choice of mode and crossing is carried out by a nested logit model, while the lower level i described in the following.

It is more difficult to generate a discrete choice set when the number of alternatives becomes larger. An essential part of route choice models is therefore to obtain efficient methods to generate and handle alternatives before to – or simultaneous with – the final route choice and assignment of trips at the more detailed level of network modelling.

The simplest approach is all-or-nothing, which is too restricted for most choice combinations. Extending this to a k-shortest path algorithm does not solve the problem on larger network, since this will often lead to a number of variations of one main route, and may overlook other main alternatives (unless 'k' is very large).

A commonly used approach is to use a set of deterministic rules to generate the choice set, e.g. fastest, shortest or cheapest route. However, it is not certain that this will include all relevant alternatives for the following choice model, e.g. if the optimal route is neither the fastest nor cheapest. The entire solution space of combinations of coefficients in a deterministic utility function may also be search, which for a limited number of variables turned out to be efficient (Nielsen & Jovicic, 2003b). This is however an infeasible approach if the utility function has more than 4 variables.

Discrete choice-sets may accordingly not be complete for larger network sizes. To avoid this, the modelling frameworks in the Copenhagen region therefore use a method that simultaneously generates paths and assign traffic to them. Different methods for this have been proposed in the literature;

1. Dial (1971) developed an algorithm to solve a logit-based assignment in a network. The generation of the full choice set and path probabilities are very efficient. It is however assumed that all routes are uncorrelated (the general assumption behind the logit model). This is however doubtful in most real network (Sheffi, 1985).

2. Dial (1997) formulated a so-called bicriterion algorithm, which makes it possible to solve a route choice model with both time and cost in the choice function, where the value of time (VoT) follows a distribution. The model does however not include an error term (does not consider the general overlapping route problem) and is restricted to two variables only (although each can be the sum of several components).

3. The third approach is to use simulation to simultaneously generate choice sets and route probabilities by an approximation to the probit model. The difference to the logit model family is that the error term is simultaneous normal distributed (opposite independent Gumbel) over alternatives, and that the variation of the error term is proportional to the mean utility of the alternative (opposite identical Gumbel). The principle in the solution algorithm (Sheffi, 1985) is that the error term is simulated in an inner loop while all-or-nothing routes are generated. An outer loop (the Method Successive Averages, MSA) averages the results with prior iterations. The distribution between routes

approximates a probit model as the number of iterations increases. The calculation complexity is identical with the complexity of generating a choice set to a discrete choice model by simulation, and by using MSA it is not needed to store paths (which can be infeasible for larger networks due to computer RAM restrictions).

Although the third approach is fairly calculation demanding, it has been used in several large-scale models in Denmark with success, e.g. in the Harbour Tunnel model for cars, vans and trucks route choice (Nielsen et. al., 2002), frequency-based assignment in Nielsen (2000) and timetable-based in Nielsen (2004). Since the model type can combine both an error term, random coefficients in the utility function and flow-cost relationships, it has been chosen for the freight assignment in the present paper. As freight network assignment is usually done at a more aggregated level of detail (e.g. that fewer lines exist than in regional public transport), the calculation time and convergence becomes a less critical issue than when used for passenger transport. The model is however more difficult to estimate and calibrate due to the complex utility functions. For this reasons the Danish freight model prestudy (Nielsen et.al. 2003) proposed a hierarchical model structure with MNL at the upper level, and only use of the probit model at the lowest level.

4 Equilibrium principles

Wardrop's (1952) first principle defines the *user equilibrium* in a traffic network as the state in which no passenger can reduce his/her travelling time solely by changing route. Daganzo & Sheffi (1977) extended this principle to the Stochastic User Equilibrium (SUE):

> *An equilibrium is obtained where no traveller's perceived cost can be reduced solely by the traveller changing route.*

The *perceived* corresponds to the fact that travellers may not have complete knowledge of the network, and to other random variation. The *cost* is a generalisation of only one variable, namely time consumption. This can be reformulated to a utility maximisation with any *utility function*, where the sign is reversed. Daganzo & Sheffi (1977) formulated SUE as a principle and mathematical programme, and Sheffi & Powell (1982) found an operational solution algorithm hereto (Method of Successive Averages, MSA). The East Denmark Model (EDM, Nielsen et.al. 2001a) used an extended version of the user equilibrium:

> *An equilibrium is obtained where no traveller's perceived utility, determined by the traveller class' utility function, the traveller's preferences, and the type of vehicle*

> *(service) and its reliability can be increased solely by the traveller changing route at the desired time of travel.*

This implies an equilibrium with a discrete number of traffic classes, each with its own utility function, variation of its coefficients (variation of preferences within each class modelled as random coefficients) and error term (unexplained variation, handling of overlapping routes). The vehicle type (service) influences reliability as well as the Level of Service (LoS) within the utility function. The utility on a link is allowed to influence other links, e.g. if a full ship in a given departure prevents adding more containers on it. Finally, the LoS depend on the desired time of travel through the timetable, punctuality (delay distributions) and capacity. The above equilibrium principle is slightly modified to consider freight:

> *An equilibrium is obtained where no freight categories' generalised cost function determined by the freight class' generalised cost function, the transports preferences, the type of vehicle and its reliability can be increased solely by the freight category changing route at the desired time of travel.*

Most assignment models assume increasing travel costs with volume. However, this seems improper at the level of route segments choices in freight transport;

- Travel times may reduce with volume, due to higher frequencies and maybe use of more efficient means of transport.

- Travel costs will in most cases reduce with volume (demand), since larger more efficient means of transport can be used, and they can be used more efficiently as volume increase. Examples are container ships, which have low unit prizes on high demand routes, and truck transport that is cheaper between regions with high demand and competition than between other destinations.

Capacity restrictions (a decreasing speed flow curve) are then mainly relevant for the transport network as one would normally expect that the market will provide sufficient transport capacity in terms of means of transport.

5 Utility functions for freight transport assignment

The utility function for a given transport is described by the class of transport it belongs to. High value goods for just in time production has e.g. higher values of time and demands to precision compared to low value heavy goods. The preferences (coefficients) within each class may also differ, which is modelled by random coefficients.

The line-based transport network refer to the physical network by using an object oriented GIS-model similar to Nielsen & Frederiksen (2001b). The lines may refer to the same links that are used for individual traffic assignment, e.g. that line-based trucks use the same road as door-to-door trucks. Or lines may run on its own network (typically rail).

Each mode *i* is described by specific coefficients on the travel time, as well as possible other attributes X_{ji}, where *j* indicates the vector of attributes. Random coefficients are included in the utility function to consider different preferences within groups. These are added as stochastic terms ξ_j to the coefficients β_j. It is often assumed that the distribution of one coefficient is uncorrelated with the others. However, one may assume that a high value of one coefficient *j* (e.g. free flow time) is correlated with another coefficient *j'* (e.g. congestion time). The general framework therefore allow for a matrix of correlation, i.e. *j x j'*;

$$U_i = \sum_j \left(\beta_j + \xi_{j\,x\,j'}\right) X_{ji} + \varepsilon_i , \qquad (1)$$

Where the error term ε_i (Gumbel distributed) can be generalised to allow for a nested logit model formulation to employ e.g. different Stated Preference (SP) datasets used for the estimation.

The choice context includes many alternatives with the same main mode. The choices are accordingly not only between a discrete set, but rather due to different preferences with regards to attributes within a set of combined choices. The random coefficients describe different weights (priorities) between attributes, while the error terms describe choices between main alternatives (e.g. binary choices in a Stated Preference experiment used to estimate the model).

The choice situation in real schedule-based networks includes varying knowledge on the network. Arc-based variation, $\varepsilon_{a,}$, are accordingly added to describe the choices in the network;

$$U_R = \sum_j \left(\beta_j + \xi_{j\,x\,j'}\right) \sum_{a \in R} X_{ja} + \varepsilon_R + \sum_{a \in R} \varepsilon_a , \qquad (2)$$

Where *a* indicates the vector of arcs, e.g. along a route *R* between a zone pair. Note that the choice-set consists of routes *R* instead of alternatives, *i*.

6 On the random coefficients, ξ

Importance of heterogeneous preferences is widely recognised in passenger transport models. Freight or firms do not have preferences or utilities like people do. Still it is very likely that heterogeneity is just as important for freight. This can be described by random parameters in the generalised cost functions. In general, the Danish pre-study (Nielsen et.al. 2003a) recommended to model as much heterogeneity as possible through a parameterisation of the general cost function. Nevertheless, there might be enough remaining heterogeneity that explicit modelling of this is needed, due to the following reasons:

- In general large heterogeneities can be found even within a sector (e.g. value of goods). Transport of fresh fish for sushi will e.g. most likely have a higher value of time than fish remains to be used in the animal food industry, although both are within the same (statistical) goods categories.

- There might be variation of preferences within a certain category of goods, e.g. due to different business strategies and concepts.

- There might be variation of time restrictions due to different production principles, e.g. just-in-time production versus in-house warehouses, in-house production versus outsourcing, etc.

Although most random coefficient models in practice use normal distributions, this is not done in the present model for the following reasons;

1. The normal distribution provides also negative values, which is in conflict with the assumption that e.g. time use represents a cost.

2. The negative values cannot be part of route choice models; since this will increase the algorithmic complexity considerably (Dijkstra can e.g. not be used).

3. The Value of Times (VoT) are undefined with normal distributed cost coefficients (if the VoT is calculated as a normal distributed time coefficient divided by a normal distributed cost coefficients, this follows a Gauchy distribution which is undefined in mean and variance).

An often used alternative to the normal is the logarithmic normal distribution, which has a fairly uncomplicated functional form. It is multiplicative which means that the VoT is also log.normal if the cost and time coefficients are log.normal. The work on Danish data indicates that log.normal distributions seem to be the best choice in most cases, and that especially the distributions of time-coefficients are highly correlated. The log.normal distribution has also the advantage of being non-negative. Nielsen et.al. (2002) included distributed coefficients for vans and trucks, which is adopted for the present model.

7 On the error term, ε

In some circumstances it is a myth that freight transport act more rational in term of generalised transport cost than passenger transport, e.g. due to;

- The cost of knowledge acquisition (time use). If a logistic manager earns a high hourly wage, it is not rational to use a lot of time to reduce transport cost for infrequent transport tasks.
- Transport costs are often infinitesimal compared to production cost, capital cost, ware house costs, etc. Therefore transport costs may not be a core area of focus in many companies who buy transport.
- Transport buyers use often the same transport company due to habits or convenience, standard contracts (i.e. fixed partner of co-operation) or local preferences (this is e.g. fairly frequent in some provincial regions).
- Some companies have their own transport departments that are used for all transport (only the marginal cost is considered for each transport).
- Each transport company may not have access to the entire network (e.g. not to some terminals and transferring points), or they may travel in some parts of the network very infrequent.

For these reasons it can be the case, that the error term has as large or larger variance compared to the mean utility than for passenger cars. It is to be preferred that the arc-based variation for this follows additive distributions in order to make the route choices independent of the specific digitisation and segmentation of the network. This is secured by

1. Forcing the variance of ε_a to be proportional to the mean of the utility, i.e. $\mathrm{E}(\cdot) = \rho \cdot \mathrm{var}(\cdot)$;
2. Using a distribution which is additive in mean and variance (van Vuren, 1995, and Nielsen, 1997).

Using a normal distribution fulfil this in the case without truncation (Sheffi, 1985). But the condition is violated if truncation is likely. The probability of truncation depends on ρ as well as the mean utility, since the probability increase with decreasing mean utilities. An alternative is the Uniform distribution that approximates the normal distribution on large networks with many iterations due to the central limit theorem. This works better than the truncated Normal (Nielsen, 1997), but it does not eliminated the problem with truncation fully, and posses a potential convergence problem. The Gamma distribution fulfils the two conditions directly, since it is additive and positive, whereby truncation is avoided (Nielsen, 2004).

The overlapping route problem must be solved for fixed coefficients β, i.e. on a given homogeneous user segment. When solving the equilibrium model in practice, this imply that the simulation of the stochastic coefficients must be carried out prior to the simulation of ε_a within each iteration in the solution algorithm. When *applying* the model, ρ must be calibrated by simulation and compared with observed route choices. Accordingly, the estimation and calibration procedure can be formulated as:

- Estimate utility functions (coefficients and random coefficients) as in formula (1);
- Transfer these to the assignment model and calibrate ρ by simulation (this is an uncomplicated task as the problem is one-dimensional);
- Carry out convergence tests for the selected ρ to determine the needed number of iterations in the assignment procedure, since this may depend on the level of variation in the model.

8 The SOLUTION algorithm for schedule-based transport

The solution algorithm for schedule-based transport build on the work presented in Nielsen (2004), as the problem is quite similar for passenger and freight transport (except that different networks, utility functions and rules are used). As in passenger transport, feeder mode links may not be schedule-based. Trucks may be considered as a feeder mode in freight assignment to rail and ships, while trucks themselves can operate according to schedules as well (e.g. for parcel post and mixed cargo). Non-schedule-based freight trucks (and vans) can be modelled as in Nielsen et.al. (2002), where they are assigned simultaneous with passenger cars.

The upper level schedule-based problem is typically solved at the route segment level (Nielsen & Frederiksen, 2001b), i.e. where a route segment connects to two terminals can be considered as an arc. At the lower network modelling level, the traffic is preloaded at the links and nodes along the segment in the graph before the remaining non-schedule-based transport is assigned. Nielsen (2004) found three possible ways to optimise such utility-based stochastic timetable-based models:

1. Optimisation of the implementation, especially building the graph a priori by a pointer structure in memory speeded up calculation time. This included pointes for all heaps and search trees.

2. Optimisation of the shortest path algorithms by use of other methods than the classic Dijkstra, e.g. iterative algorithms, by reducing the size of the calculation graph (removal of irrelevant parts of the graph in a way effective for calculation), and by using the principle of event dominance

3. Optimisation of the solution algorithms for the route choice model (MSA), by trimming the updating mechanism.

9 Conclusions and recommendations

The paper presents the methodological considerations behind the ongoing work on the Danish freight transport model. It was generally agreed (Nielsen et.al. 2003) that the choice problem of mode and route should be formulated as a three level problem, i.e. decision made by

1. Transport buyers, i.e. choice of which main mode combinations to consider.

2. Transport suppliers, i.e. which combinations of route segments to consider.

3. Transport routes within each route segment.

In the Danish case it was possible to describe most of 2) as an a priori discrete choice set, since the sea sounds with ferry lines and bridges between East Denmark (Zealand) and West Denmark (Jutland), Southern Sweden and Germany constitutes natural barriers. Level 1) and 2) can be dealt with by traditional discrete choice models.

The main focus of the paper was accordingly level 3). It is argued that all-or-nothing as well as deterministic equilibrium methods are too simplified, since heterogeneities in preferences within transport classes (random coefficients) and stochasticity (link-based variation) exists among transport buyers and suppliers. Each class of transport (mode and good specific) must have its own utility function of e.g. length, cost, time and congestion time, for which stochastic terms are included. Software and methods have now been developed that can estimate complicated utility functions. It is recommended to use correlated log.normal coefficients in these, since empirical evidence justify this and it is preferable to use a non-negative distribution.

An user equilibrium is much more reasonable to assume for the freight sector than a system equilibrium, and cost-volume relationships at level 2) may be lower cost, higher frequency and faster time with increased volume (due to larger and more efficient means of transport, less empty driving and more competition). This is opposite most existing models, who assume increasing cost with volume.

The solution algorithm is the Method of Successive Averages, where the stochastic parts of the utility function are simulated. Different suggestion to focus the stochastic simulation and updating mechanism in MSA to those paths being used in the present iteration can reduce the calculation time. The memory need and storage space can hereby be met by a standard PC even on a large network, why the method is fully feasible for real size models.

References

C. F. Daganzo and Y. Sheffi, “On Stochastic Models of Traffic Assignment”, *Transportation Science* 11, 253-274 (1977).

R. B. Dial, “A probabilistic Multipath Traffic Assignment Algorithm which obviates Path Enumeration”, *Transportation Research*. 5, 81-111 (1971).

R. B. Dial, “Bicriterion traffic assignment: efficient algorithms plus examples”. *Transportation Research, Part B (Methodological)* 31B, 357-79 (1997).

O. A. Nielsen, “On the distribution of the stochastic component in SUE traffic assignment models”, *25th European Transport Forum (PTRC Annual meeting)*, Proceedings. Seminar F, Transportation Planning Methods, Vol. II, 77-94. Uxbridge, UK, 1997.

O. A. Nielsen, “A Stochastic Transit Assignment Model Considering Differences in Passengers Utility Functions”, *Transportation Research Part B Methodological* Vol. 34B, Elsevier Science Ltd., 337-402 (2000).

O. A. Nielsen and C. O. Hansen and A. Daly, “A large-scale model system for the Copenhagen-Ringsted railway project”, *Travel behaviour Research: The Leading Edge*, David Hensher (edt.), Pergamon press, Chapter 35, 603-623 (2002).

O. A. Nielsen and R. D. Frederiksen, ”Rule-based, object-oriented modelling of public transport systems – A description of the Transportation Object Platform”, *9th World Conference on Transportation Research (WCTR),* Pre-prints, session D1-01, 23/7. Seoul, Korea. A new version is to be published in the selected proceedings, Elsevier, (2001).

O. A. Nielsen and R. D. Frederiksen, ”Optimising Timetable-Based Stochastic Transit Assignment Models”, *Triennial Symposium on Transportation Analysis (TRISTAN IV).* São Miguel, Zores, June. Preprints, Vol. 1/3, 195-200. Forthcomming in Annals of operations research, (2004).

O. A. Nielsen and R. D. Frederiksen and A. Daly, “A stochastic multi-class road assignment model with distributed time and cost coefficients”. *Networks and spatial economics* 2, Kluwer, 327-346 (2002).

O. A. Nielsen and M. Fosgerau and C. O. Hansen and M. Holmblad and J. H. Rich, “*A national freight transport model – recommendations*”, (written in Danish), Report, Danish Transport Research, (2003a)

O. A. Nielsen and G. Jovicic, “The AKTA road pricing experiment in Copenhagen”, *10th International Conference on Travel Behaviour Research.* Proceedings, session 3.2 Valuation/Pricing. Lucerne, Switzerland, August 2003.

O. A. Nielsen, “A large scale stochastic multi-class schedule-based transit model with random coefficients”, *Schedule-Based Dynamic Transit Modelling – Theory and Applications.* Chapter 4, Nigel Wilson and Agostino Nuzzolo (eds), Kluwer Academic, 51-77 (2004).

Y. Sheffi and W. B. Powell, “An Algorithm, for the Equilibrium Assignment Problem with Random Link Times”, *Networks* 12, 191-207 (1982).

Y. Sheffi, “Urban Transportation Networks”, Prentice Hall, Inc, Englewood Cliffs, NJ, (1985).

T. Van Vuren, “The trouble with SUE stochastic assignment options in practice”, *PTRC Summer Annual Meeting***,** Proceedings, Seminar H. 41-52. University of Warwick, England, 1995.

J. G. Wardrop, “Some theoretical aspects of road traffic research”, *Institution of Civil Engineers*, London, (1952).