C O M P A S -
Copenhagen Model for Passenger Activity Scheduling

Goran Jovicic, civil engineer (gj@dtf.dk)
Danish Transport Research Institute (www.dtf.dk)
Department for Transport Economics and Modelling

1. Introduction

The focus of the paper resides in a description of the proposed structure for a travel demand model for Copenhagen (section 3). The model is named COMPAS, which is an acronym for Copenhagen Model for Passenger Activity Scheduling.

COMPAS belongs to the family of activity based demand models that are theoretically superior to the conventional trip based demand models. These two groups of models are compared in a separate section of the article (section 2).

2. Background of the activity based travel demand models

Activity based travel demand models predict travel behaviour as a derivative of activities. This means that by predicting what activities are performed in a day each trip to the activity is modelled through the choice of destination, choice of time of travel and mode choice in activity based models.

An activity can be defined as a physical engagement of an individual in something that satisfies his or family needs. Activities are motivated by economical, physiological and sociological needs of an individual. Activities can be grouped into various categories, e.g. work, shop, recreation, mandatory, optional, etc. An activity does not necessarily result in a trip, i.e. many activities are completed at home. A decision to engage in an activity represents a complex interaction of:

- Household and individual roles and responsibilities.
- A particular lifestyle of an individual and his family.
- Options on activity type, location and duration.
- Time, space and budget constraints.

Let us imagine that a person has completed following activities in a hypothetical day (figure 1): He went to work in the morning, then he went for a meeting in the midday and later he returned back to work, then he went home and on the way he shopped in a supermarket. In the evening he went to the cinema and afterwards he went home. Four activities (work, meeting, shopping and leisure (cinema) were therefore completed with seven accompanying trips. He also completed two home-based tours (one with an additional stop to the supermarket, and the other without stops) and one work-based tour (without stops).

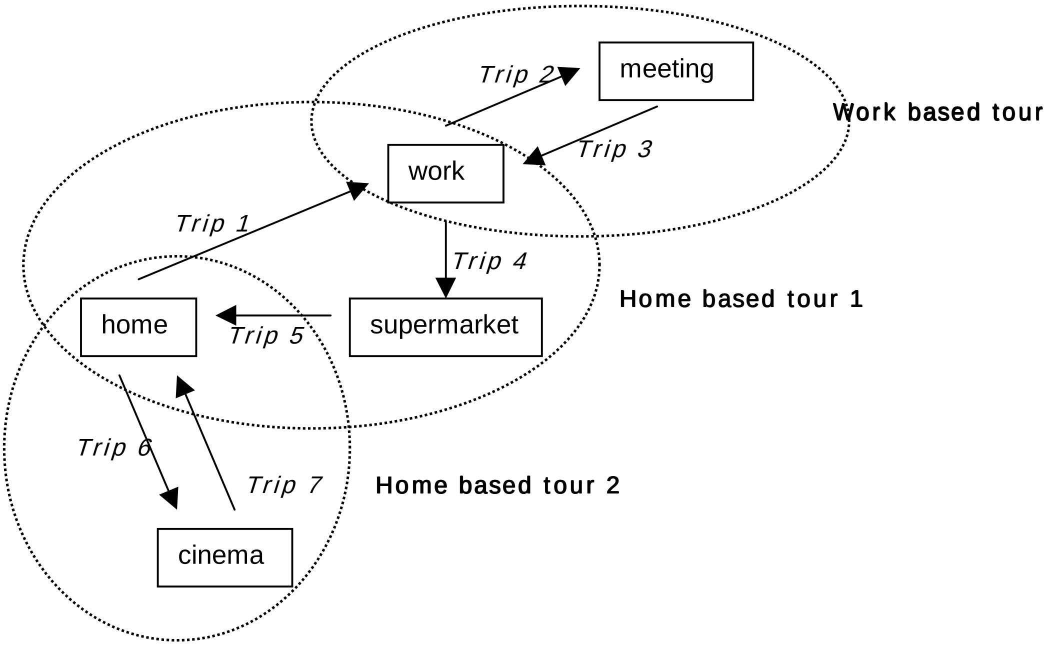

Figure 1: A graphic presentation of activities and trips completed in a hypothetical working day by an individual

Trip based demand models would model the seven trips independently of each other. Tour based demand models would model tour 1 (home-work-home) and tour 2 (home-cinema-home) independently of each other, while the work based sub-tour (work-meeting-work) would be modelled as two independent trips. Activity based models would, on the other hand, model the four observed activities and the observed trips as parts of the same decision process. Therefore, all interdependencies between the completed trips are captured in this kind of models.

If we now introduce, say, road pricing in the rush hours only, the person can react in a number of ways. He can, for instance, decide to stay and work at home that day, change the departure time for specific activities in order to avoid road pricing and/or choose different travel modes or destinations for some activities. Personal constraints and obligations (roles) will obviously contribute to a very complex set of changes caused by road pricing. The above listed changes cannot be captured entirely in the trip based and tour based demand models because these models cannot recognise the entire complexity of the newly existing situation.

The following five important features of the activity based paradigm are:

- Travel is a derived demand from the activity participation.
- The activity based approach focuses on sequences of patterns of activities.
- Individuals' activities are both planned and executed in the household (family) context.
- Activities are spread through out a 24-hour period in a continuos manner, rather than using the simple categorisation of 'peak' and 'off peak' events.
- Travel and location choices are limited in time and space, and by personal constraints.

Taking the above into consideration, it can be argued that the activity based approach to travel demand modelling gives a rich and accurate framework in which travel is analysed as a daily pattern of behaviour, related to and derived from differences in lifestyles and the activity participation among individuals.

There are two main streams within the activity based modelling of travel demand, i.e. discrete choice models and rule based simulation models. The essential difference between the two approaches is that *the time component is modelled discretionary in the discrete choice models while in the simulation models the time component is modelled continuously*. Details of the two model groups are described in Jovicic, 2001.

2.1 Discrete choice activity based models

Discrete choice activity based models are based on the random utility theory. Structurally, they represent a qualitative extension of discrete choice trip based models. The extension contains improvements in the traditional generation model where now activities (and their related trips) of different purposes are combined in what is called an 'activity pattern' model. Secondly, discrete choice activity based models chain trips into a day-overall activity pattern, taking care of the existing constraints. These models can be pictured as large nested logit models where on the higher level an individual chooses a travel pattern and below that are placed trips associated with these activities. The alternatives (activities, destinations, modes of travel and times of travel) are described through utility functions. The model levels are connected from the bottom to the top via accessibility variables. In that way the performance of transport services (that are placed on the trip level) have also impact on the activity patters. Prior to day travel patterns, these models usually model some aspects of the life-style such as the choice of work location and the choice of car ownership. Important strengths of discrete choice activity based models are the following:

- A large set of activities is defined in the travel pattern model. They are defined upon the type of the main activity, primary tour structure, secondary tour structure, pattern of intermediate stops, etc.

- Long term effects are included in the model structure.

- Attributes of the transport system performance are included in the model structure.

- Discrete choice models are based on the random utility theory, which leads further to the probability models. This is a more accurate approach to modelling choice behaviour than the approach assuming complete consistency in the way people perceive and express their preferences.

- Statistical validation of the model estimates is possible through the application of commercial software.

- Forecasting results of discrete choice models can be validated in a number of ways.

With this approach of modelling travel demand, discrete choice activity based models have been often applied when building operative traffic models that are aimed for long-term forecasts. The Portland model (Bowman, 1998) is the most famous model in this family of models.

2.2 Simulation activity based models

Simulation activity based models apply some kind of learning mechanism in order to explain how individuals build activity schedules. The idea with the learning mechanism is that we seldom consider all available alternatives when planning activity schedules. According to the theory behind simulation models, we tend to base activity decisions on some heuristic rules that are applied at some specific situations that we recognise from the past. If a new situation occurs we will again tend to consider only those activities that seem to be logical or appropriate to us. Usually, based on an externally existing activity program the patterns are built step-by-step along the day in these models taking care of the accidental circumstances that can influence the planned activities. Time is modelled continuously in simulation models. Finally, time constraints, location constraints and budget constraints are incorporated explicitly in these models.

Simulation models are often based on specific analyses such as stated preference analyses. With this approach of modelling travel demand, simulation models are mostly applicable for specific planning tasks for which short-term forecasts are needed. A Dutch developed ALBATROSS model (Artenze, 2000) is probably the most famous model in this family of activity based travel demand models.

3. COMPAS

3.1 Model description

Copenhagen Model for Passenger Activity Scheduling, COMPAS, is meant to be an operational travel demand model for the Greater Copenhagen Area, GCA. The desire to develop an activity based travel demand model for the GCA is based on the theoretical superiority of this type of models relative to the conventional trip based models, such as the Orestad Traffic Model (OTM) (Jovicic and Hansen, 2002) and the København-Ringsted Model (Nielsen, 2000).

COMPAS models activity scheduling of a person in four levels (figure 2):

- **Level 1: Modelling of mobility and lifestyle.** Two models are included here. These are a car ownership model and a model for work destination.

- **Level 2: Modelling of the activity pattern.** A model describing the number of activities, their types and order of execution of activities performed by a person in a day is included here. (See section 3.2 for more details regarding this model.)

- **Level 3: Tour based modelling.** For each home based work, school and leisure tours three types of outcomes are modelled at this level. These are the tour's time-of-day, destination and travel mode. The leisure tours are divided in several categories, e.g. shopping, visiting family/friends.

- **Level 4: Trip based modelling.** In the case that intermediate stops are made when travelling to the main destination of a tour, it is necessary to model departure time, destination choice and the mode choice. These models are conditioned to the modelled departure time, destination choice and the mode choice from level 3.

Models in the first level are understood as medium-term and long-term models, while the models in levels 2, 3 and 4 are day models. The models on lower levels are conditioned to the models on the higher levels, e.g. the trip based mode choice model is conditioned to the chosen mode for the tour.

In the opposite direction, the higher level models are influenced by the lower level models through the measure of accessibility, i.e. logsums. For instance, if a person needs to make an intermediate stop to deliver a child to the school on the way to work, because of what car is preferred, the car-mode may be also preferred for the home-work tour. In a similar manner, an improvement in the public transport service will be reflected in the person's activity pattern (e.g., an extra shopping tour might be completed) due to inclusion of logsums in utilities of the activity pattern model.

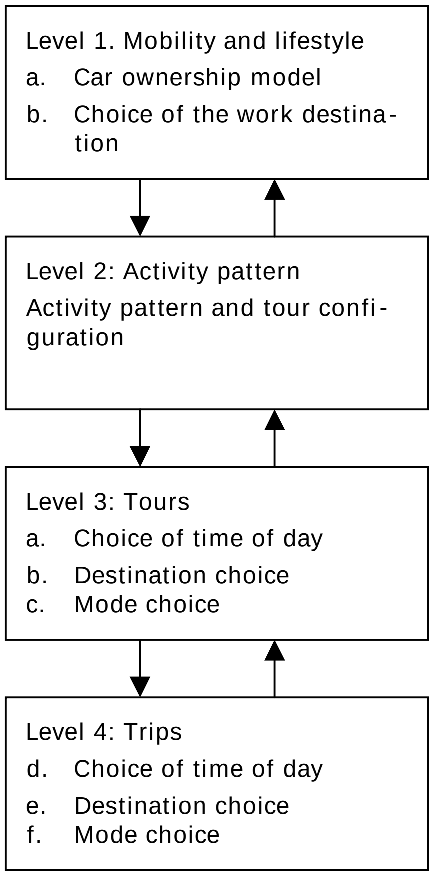

Figure 2 – The basic structure of COMPAS

COMPAS belongs to the group of discrete choice activity based models because of what it will be developed upon the experiences from the existing Danish large scale models, such as the OTM. The model operates on a 24-hour base, for an average working day.

3.2 Day activity pattern model

Theoretically most interesting model in COMPAS is the day activity pattern, DAP, model. The DAP model is a multinomial logit model where the alternatives are defined by:

- The choice of the primary activity of the day (see below for details).

- Whether the primary activity occurs at home or on the tour.

- Type of tour for the primary activity, including the participation and purpose of any intermediate stops before or after the primary stop. One such a tour can include a home-work activity as the primary activity of the day, where the person is also involved in a shopping activity on the way home from work.

- For work patterns, the participation and purpose of a work based subtour (e.g., going for a meeting) defined also the tour type.

- The number and purpose of secondary tours. A secondary tour can be a tour to the cinema in the evening, after a home-work tour was completed earlier in the day.

The primary day activity is classified in the following groups in COMPAS:

- Subsistence (work/school) at home.

- Maintenance (personal business) at home.

- Discretionary (social, recreational) at home.

- Subsistence (work/school) on tour.

- Shopping on tour.

- Other personal business on tour.

- Discretionary (social, recreational) on tour.

If the primary activity is completed on tour, the tour configuration is defined in the activity pattern model. There are defined four types of tour patterns in COMPAS:

- A simple pattern, i.e. that is without stops between home and the destination place.

- One or more intermediate activities on the way from home to the primary destination.

- One or more intermediate activities on the way from the primary destination to home.

- One or more intermediate activities in both directions.

The model allows individuals to make an extra activity (tour) when they are at work. This is defined to be a work-based sub-tour.

Once when the primary tour is finished an individual can make a secondary activity (tour). The model defines four types of secondary tours defined based on the number and purpose of these tours:

- No secondary tour.

- One or more secondary tour for work or personal business.

- One or more secondary tour for leisure.

- Two or more secondary tours.

Such a complex structure of the DAP model can result in a logit model with 1.000-1.500 alternatives (utility functions). The DAP model is based on the socio-economic characteristics of individuals (TU data), zonal data, LOS files and logsum variables.

3.3 Data requirements

COMPAS is based on variety of data:

- Description of the conducted activities and the resulted trips,

- Respondents' socio-economic characteristics,

- Zonal characteristics, and

- Network description.

The model applies the national travel behaviour survey data, TU data, for the period 1997-2001. Some 1.500 interviews are completed every month, where some 30% belong to the GCA. On the annual base we are dealing with some 5.000 respondents in the analysed area. The total sample in the model is therefore based on some 25.000 respondents and it is assumed here that this sample is a representative sample of the GCA population.

The presently available information in the TU data falls into two categories, i.e. information related to the conducted trips (activities) and the socio-economic data for the interviewed person and some household information.

3.4 Model outputs
COMPAS is a disaggregate model both in the estimation and forecasts, i.e. the model predicts activity schedules for each individual in the TU sample applying the methodology of synthetic population (see the appendix for more details regarding the methodology).

COMPAS produces three types of results:

1. The basic output of the model is an activity pattern for each individual in the TU sample.

2. Aggregation of results for policy analysis. That is, the individual outputs can be aggregated according to socio-economic characteristics for which reliable variables are available at the individual level (e.g., age, gender, income). For instance, we can measure here whether road pricing has a greater impact on travel demand for low-income families compared to higher-income families.

3. The individual activity (trip) patterns can be aggregated on the zonal level, which produces origin-destination trip matrices. The trip matrices are split by mode, trip purpose and time-of-day, and they can be assigned on the road and public transport networks. The assignment runs must be performed outside COMPAS.

References

Artenze, T. et al. (2000). 'Rule based versus utility maximising models of activity travel pattern', 9th International Association for Travel Behaviour Research Conference, Australia

Bowman, John (1998). 'The day activity schedule approach to travel demnad analysis', Ph.D. Thesis at the Massachusetts Institute of Technology, USA

Jovicic, Goran (2001). 'Activity based travel demand modelling – a literature study', The Danish Transport Research Institute, Note 8

Jovicic, Goran and Christian Overgaard Hansen (2002). 'A passenger travel demand model for Copenhagen', forthcoming in Transportation Research, Part A

Nielsen, Otto Anker, Andrew Daly and Christian Overgaard Hansen (2000), 'A large-scale system for Copenhagen-Ringsted railway project', 9th International Association for Travel Behaviour Research Conference, Australia

Appendix

Input to the methodology of synthetic population are following information found in the TU data: household size and structure, information regarding relationship between the household members, household income, number of cars in the household, location of the residence, employment status of household members, age and gender. (The activity behaviour in the TU sample is not important.)

The GCA population and the TU sample will be divided in a number of groups (cells) according to age-classes of the head of the household, household sizes and car ownership. There will be defined up to 48 cells (i.e., 4x4x3 classes, for the three groups respectively) in COMPAS.

For the forecasting year, the sizes of 48 cells in the sample are calculated based of the exogenous indicators. The population sizes of the 48 cells in the forecasting year are than calculated based on the base year population cell-sizes and the cell-sizes in the forecasting sample.

If, say, sell number 5 of zone 10 has 13 individuals (that is in the population) in the forecasting year then 13 respondents will be drawn from the corresponding cell in the forecasting sample applying Monte Carlo simulations. For each of the 13 drawn individuals the activity schedule is calculated in the model in the form of the calculated probabilities based on the known characteristics of these individuals. For each modelled outcome a random number between 0 and 1 is then drawn in order to simulate a particular outcome according to the modelled probabilities. For instance, let us imagine that in the destination choice model zone 1 has a probability of 0.7 to be chosen while zone 2 has a probability of 0.3 to be chosen for the drawn individual. If 0.8 is drawn then zone 2 is chosen for the destination of that particular trip.