Validation of a Continuous Simulation Model for Daily Travel

Fabian Märki
David Charypar
Kay W. Axhausen
Validation of a Continuous Simulation Model for Daily Travel

May 2012

Abstract

This paper validates a microscopic travel demand simulation that employs a continuous planning approach with an open time horizon. It uses behavioral targets to model people’s motivation to execute activities. People’s behavior originates from a planning heuristic making on the fly decisions about upcoming activities. The planning heuristic bases its decisions on the available activity execution options in the near planning future, its current execution effectiveness and on a discomfort measure derived from deviations between people’s performance and their behavioral targets. We validate the model and illustrate its features through three model configurations and suggest directions for future research.

Keywords

validation of target-based model, continuous activity generation and scheduling, microscopic travel demand simulation
1 Introduction

Microscopic travel demand simulation softwares simulate virtual people (referred to as agents) individually. For instance, Balmer (2007) uses agents which choose between different daily schedules. Activities of these schedules are executed and simulation results are handed back to the planning process, allowing agents to improve their schedules based on improved estimates of their generalized costs. This replanning step is repeated until the simulation reaches a stochastic user equilibrium with consistent travel demand and travel cost (Nagel and Flötteröd (2009)). Simulating agents individually leads to a high computational complexity which often results in computational performance and memory issues. Microscopic models typically introduce restrictive constraints to counter such issues. For instance, Balmer limits the maximum simulation horizon of standard size scenarios to a single day, making it difficult to investigate effects occurring over a period of days or weeks. Another limitation is that agents must commit themselves to a specific day-plan, making it challenging to simulate unexpected events realistically (Charypar et al. (2009) and Dobler et al. (2012)). As a consequence, a different simulation approach becomes necessary that is capable to model demand continuously, i.e. agents should be able to make decisions about upcoming activities on the fly and with an open time horizon.

We proposed a microscopic travel demand simulation in Märki et al. (2012) that utilizes behavioral targets to represent agents’ decision space. Targets can represent social and cultural norms and are closely related to observed behavior like execution frequency and time spent for an activity. Agents continuously track their performance and compare it to their behavioral targets using observation windows of different durations. Deviations from the desired behavior cause discomfort which is conveyed to a planning heuristic, making decisions about future activities agents should execute. This enables agents to react spontaneously to unexpected events. At the same time, it also reduces memory consumption because agents do not need to keep track of complete daily schedules.

The remainder of this paper is structured in two major parts. The first part introduces the target-based model and the decision heuristic. The second part validates the target-based model and the decision heuristic by using three different model configurations which base on a six-week continuous travel diary. The first configuration (base case) calibrates the model and shows that it reproduces underlying measures without additional constraints, the second configuration (exact case) validates the decision heuristic by showing how it can be brought to execute specific activities by reproducing a complete schedule, and the third case (use case) shows how we configure the model using a person’s weekly execution pattern/rhythm and compares simulation results to the behavior observed in the data using various statistics.
2 Other Work

Arentze and Timmermans (2006) introduced need-based theory and proposed a model for activity generation (Arentze and Timmermans (2009)) that assumes utilities of activities are a dynamic function of needs. Whereas Arentze and Timmermans used needs as people’s motivation to execute activities, we see the satisfaction of needs as one possible target in our model. Generally, we assume that people describe their desired performance through measures which are closer to data found e.g. in Swiss Federal Statistical Office (2006). We pick up Gliebe and Kim’s (Gliebe and Kim (2010)) suggestion to use time-dependent utilities and introduce time-dependent effectiveness functions, describing the effectiveness of activities and locations towards discomfort reduction. We presented a need-based model in Märki et al. (2011) that was also designed for a continuous simulation. Our new model (Märki et al. (2012)) drops the need-based approach and introduces measures (we refer to them as targets) which are more related to data found e.g. at the Swiss Federal Statistical Office.

3 Introduction of Model and Decision Heuristic

Agents are the central component of our model and represent virtual people. Each agent has a motivation to execute activities and specifies its desired performance through behavioral targets. Deviations to behavioral targets result in discomfort which induces agents to take action against the deviation; higher deviations result in higher discomfort which in turn leads to a higher urge to take action. Agents can reduce discomfort through the execution of activities at different locations. We assume that agents prefer activity-location pairs that provide more discomfort reduction. This is similar to Arentze and Timmermans’ work (Arentze and Timmermans (2009)), where they proposed activity utility as a function of need reduction.

3.1 Model

3.1.1 Targets

The core assumption of this work is that people have a motivation to execute activities and that they have a perception of their motivation in form of a desired performance. People specify this performance through behavioral targets and try to comply with them across observation windows of different duration. For instance, a person would like to play $2^{0.5}$ hours of tennis
about $2^{+1}$ per week. This targeted behavior is transformed into following targets:

- The **percentage of time** target defines the time a person would like to spend for an activity within an observation window. In order to simplify modelers’ task, it is possible to specify the total execution duration and the conversion to the percentage of time target is done internally. For the above example, the modeler would specify a target value of 2 hours of tennis, a bandwidth of $^{+0.5}$ hours (upper and lower bound of the target value) and an observation windows of one week (see Fig. 1(a)).

- The **frequency** target defines the number of activity executions a person would like to accomplish within an observation window. For the above example, the modeler would specify a target value of 2 executions of tennis with a bandwidth of $^{+1}$ executions and an observation windows of one week.

Agents monitor their performance during simulation and compare these state values to target values (see Fig. 1(b)). State values are exponentially discounted over the observation window of targets. This simulates a forgetting process where agents give recent behavior more weight and gradually forget their past performance.

### 3.1.2 Effectiveness Functions

Effectiveness functions inform agents about the effectiveness of activities and locations towards discomfort reduction. This is similar to Gliebe and Kim (2010) who proposed time-dependent utilities. Effectiveness functions are a broad concept and can model different effects. Possible examples are:

- **Shop opening hours for a daily shopping activity.** Agents can use this information to either determine if they can shop and for how long or how long it takes until they can shop next time. Since effectiveness functions can be location dependent, it is also possible to model location dependent shop opening hours. Furthermore, this effectiveness function can also contain time dependent information about shop crowdedness. Hereby, we assume that shopping at overcrowded shops is less efficient (smaller value) and therefore takes longer.

- **Daylight intensity for a sleep activity.** This function specifies the light intensity. Agents can use this information e.g. as an indication of sleep effectiveness. Hereby, we assume that people sleep at night and have already adapted to their current timezone.

- **Business hours for a work activity.** This function can be seen as a cultural norm (cultures may have different business hours) and a social norm (social groups, e.g. professions, may have different business hours). Agents can use this information e.g. as an indication of
work effectiveness. Hereby, we assume that people depend on co-workers to be able to do their work (the degree can differ depending on the profession).

- **Seasonal effects for a sport activity.** This function is location dependent and combines different effects like time of the year and weather conditions. As an example, a ski resort can have a high effectiveness during the winter months after a snowfall whereas the yacht club has a high effectiveness during the summer months with sunny weather and a good breeze. This enables agents to follow seasonal rhythms because they choose to ski at the ski resort during the winter and to sail at the yacht club during the summer.
3.2 Decision Heuristic

We consider a decision heuristic as a feasible approach to overcome limitations like poor performance for large scenarios (Charypar and Nagel (2006)), high computational costs (Balmer (2007)) or inflexibilities when agents should spontaneously react to unexpected events (Kuhn-imhof and Gringmuth (2009)). Since a heuristic aims to approximate a good solution, it is also possible to use incompletely knowledge about the state of mind and plans of other agents. This is helpful since complete knowledge generally induces high computational and memory costs.

The decision heuristic we proposed (Märki et al. (2012)) combines several aspects which are derived from targets and effectiveness functions. The decision heuristic takes all promising activity-location pairs, compares their heuristic values and decides to implement the activity-location pair which yields the highest heuristic value per invested time. The heuristic value is a combination of following elements (we refer readers to Märki et al. (2012) for a detailed explanation):

- **Discomfort**: The discomfort builds on targets and is a function of the difference between the target value and the state value. Discomfort can be seen as an urge agents experience to change their current situation by executing an activity that reduces the discomfort. We assume that people have a preference for activity-location pairs that yield the highest discomfort reduction.

- **Look-Ahead Measure**: The look-ahead measure builds on effectiveness functions and is calculated through the convolution of an effectiveness function with an exponential kernel that points into the future of the simulation. This gives an indication about future execution effectiveness and hence, about the flexibility to execute an activity at a later point in time. The decision heuristic uses this parameter to postpone activities with more execution options/higher future execution effectiveness and favors other activities for current execution.

- **Current Execution Effectiveness**: The current execution effectiveness builds on effectiveness functions and is calculated through the integral of the effectiveness function between activity start and end normalized by the activity duration. This measure introduces a preference for efficient time windows, whereas efficiency is defined by whatever the effectiveness function represents (e.g. social or cultural norms). This measure also offers a location choice procedure since locations can provide different effectiveness definitions.
Figure 2: Illustration of the validation procedure. The model is configured based on a person’s 
schedule and then used to simulate six continuous weeks. Statistics are extracted from 
both schedules and compared to get an understanding of the model’s capability to 
reproduce the observed behavior.

4 Model Validation

We use an existing six-week continuous travel diary (Löchl et al. (2005) and Schönfelder (2006)) 
and validate our model by focusing on the behavior of one person. This person is a 52 years old 
male who works full time on a flexible work schedule, lives in a single household but is in a 
relationship and has regular commitments. We configure three models and distinguish between 
eight activity types (home, work, social contact, shopping daily, shopping long term, leisure 
active, leisure excursion, private business). The simulation results are validated by comparing 
different statistics which are extracted from the survey and the simulation (see Fig. 2). These 
comparisons illustrate the behavior reproduction capability of each model configuration.

4.1 Survey Data

Fig. 3 illustrates the statistics extracted from the survey.

- Fig. 3(a) illustrates activity duration, frequency and percentage of time spent per week. 
The box plots come in pairs showing observations for the survey and the simulation (for 
later comparison). The whisker bars show the minimal and maximal measured value and 
the box defines the first and third quartile. The red bar shows the mean and the green bar 
the median value.
Validation of a Continuous Simulation Model for Daily Travel

May 2012

- Fig. 4(a) illustrates the average similarity of weekdays measured by Joh’s multidimensional affinity-measurement function considering activity sequence alignment and activity timing (Joh (2004)). The table shows that the same weekdays (e.g. Thursday-Thursday) have the highest similarity (darkest color). This is followed by the similarity of working days. In comparison with weekend days, working days show the lowest similarity (lightest color).

- Fig. 4(b) illustrates the execution probability of activities at weekdays. The table shows that e.g. the person executed home at least once at each day of the survey and work between Mondays and Fridays with one exceptional execution on a Saturday.

- Fig. 4(d) illustrates the transition probability between activities. The table shows that e.g. the person tends to go home after the execution of most activities (first column with the darkest color). The execution of social contact also seems common after most activities. The person tends to go to work after he was at home (dark cell in the upper right corner) and there are occasions when he consecutively executed shopping daily or social contact activities (the cross section of shopping daily-shopping daily and social contact-social contact have a light color).

- Fig. 4(e) illustrates the start and end time distribution of work. We focus on work because it is the activity where the reproduction of the start/end distribution is most important and thus, gives an indication about the scheduling quality of the model.

- Fig. 4(f) illustrates the execution intervals of activities (waiting time between activity executions). The whisker bars show the minimal and maximal measured value and the box defines the first and third quartile. The red bar shows the mean and the green bar the median value. Execution interval also validates the scheduling quality of the model. We prioritize shopping daily and social contact as goodness indicators since these activities have a repetitive nature.

4.2 Base Case

In the first case, we use static configurations for frequency and percentage of time targets. Fig. 4 illustrates, that work is configured with an average frequency target value of 10.15 executions per week and an average percentage of time target value of 24% of the weekly available time. The remaining activities are configured in the same manner.

The simulation of six consecutive weeks with the above model configuration produces the results shown in Fig. 6(a) Fig. 7(a) and Fig. 8(a).

- Fig. 6(a) illustrates that the simulation was able to reproduce mean values. This result meets our expectations, since the target based approach builds on mean values (the target values configured according Fig. 4 define the average behavior/value a person tries to
Figure 3: Illustration of the statistics extracted from the survey (see Section 4.1 for a detailed explanation).

(a) Statistics for activity duration, frequency and percentage of time spent per week.

(b) Similarity table of Joh’s multidimensional affinity-

(c) Execution probability table for activities at week-

(d) Transition probability table between activities.

(e) Start/end time distribution of activity work

(f) Execution intervals of activities
Figure 4: Base case configuration of the activity work with static values.

(a) Configuration of the frequency target.

(b) Configuration of the percentage of time target.

achieve). Because this configuration does not define any other constraints, the simulation did not reproduce the variability observed in the survey.

- The similarity table of Fig. 7(a) illustrates that days are too similar, i.e. there are no differences in working days and weekend days. The execution probability table reveals that the agent also worked during weekends and that other activities were also executed on Sundays which were not executed in the survey.

- Fig. 8(a) shows that the agent executes work at any time of the day and interval times for activities are not reproduced either. This is the anticipated result since this configuration does not define scheduling constraints.

We adopt the calibration parameters (procentual deviation of the lower and upper bound from target values) from this configuration for the remaining two model configurations.
4.3 Exact Case

In the second case, we use a dynamic six week configurations for frequency and percentage of time targets. This configuration is extracted from the survey through a convolution of the activity patterns with an exponential kernel of one week, resulting in exponentially weighted moving averages. The consecutive six week plot of Fig. 5 illustrates the configuration of the percentage of time target for home. The remaining activities are configured in the same manner.

Our expectation is that the simulation reproduces the schedule used to configure the model. This raises the question why we configure the model like this since we could directly use the schedule in the first place. It is important for the validation of the model and the decision heuristic because it shows that both components go hand in hand and comply with the configuration. It also shows how it would be possible to bring agents to execute specific activities. This is important for:

- **Concept of Project**: The concept of Project is a mechanism that temporally influences behavioral targets and is used to model non-recurrent tasks, i.e. tasks that are extra to ordinary life (e.g. get married). Consequently, a mechanism becomes necessary that enables modelers to tell agents where and when they should execute such extra tasks (see Märki et al. (2012) for a detailed explanation of Projects).
- **Social interactions**: Social interactions (e.g. within a social network) require that agents can meet at specific places and times. Again, this necessitates a mechanism that forces agents to comply with such appointments.
The simulation of six consecutive weeks with the above model configuration produces the results shown in Fig. 6(b), Fig. 7(b) and and Fig. 8(b).

- Fig. 6(b) illustrates that the simulation was able to almost exactly reproduce the survey. Slight differences in the minimal and maximal values are due to dynamic travel times, i.e. the agent misjudged the time to travel to the next location.
- Fig. 7(b) and Fig. 8(b) confirm above observations. The tables and plots are identical, showing that the simulation reproduced the survey.

The investigations of this section illustrate how the model facilitates the possibility to tell agents what activity they should execute. At the same time, it also shows that the heuristic can handle dynamic travel times. However, this is not the way we intend to use the model. The configuration procedure we intend to use is covered in the next section.

4.4 Use Case

In the third case, we use the six weeks of exponentially weighted moving averages from case two (see Section 4.3) and combine them into a weekly pattern (see Fig. 5). The idea of this procedure is to get an average behavioral pattern for the target value representing the observed weekly rhythm, and a lower and upper bound specifying the allowed deviation from the behavioral pattern. The remaining activities are configured in the same manner.

The simulation of six consecutive weeks with the above model configuration produces the results shown in Fig. 6(c), Fig. 7(c) and Fig. 8(c).

- Fig. 6(c) illustrates that the simulation results are inbetween the observations of case one and two. For instance, the simulation almost exactly reproduces the duration distribution of home, whereas the duration of work has an extreme outlier (in one occasion the agent worked for almost 23 hours). Leisure active on the other hand has a systematic deviation from the observed frequency.
- The similarity table of Fig. 7(b) illustrates that the same weekdays (e.g. Friday-Friday) have a higher similarity (darkest color) than they have to other days. Working days on the other hand have a higher similarity to other working days than to weekend days (weekend days have the lightest color). The execution probability table explains why weekend days are dissimilar to working days. The agent executes work only on working days and its execution pattern on weekend days is very similar to the observations in the survey. The execution probability table also shows that there was a Wednesday when the agent did not go home (due to the day when it worked almost 23 hours) and that it executes too
many leisure active activities (the row has a darker color than the row of the survey). The transition probability table illustrates that the agent usually also went home after most other activities and that it usually also went to work after it was at home. In the simulation, there were occasions when the agent consecutively executed shopping daily or social contact. This is similar to the observations in the survey. On the other hand, the color of the column for leisure active is too dark, showing that there are too many transitions to that activity (due to too many executions).

- Fig. 8(c) shows that the simulated and the observed start/end time distribution pattern have similar characteristics. Activity start peaks are better reproduced than activity end peaks. We consider the simulated pattern as reasonable since morning and after lunch peaks are usually more prominent than those in the evening. Furthermore, we are satisfied with this pattern because it also indicates that the model does not overfit the data and still reproduces major characteristics. The execution interval of some activities (e.g. shopping daily and social contact - both have a repetitive nature) are reproduced whereas others (e.g. leisure active) do not show enough variability. However, results for leisure active indicates that there is a flaw and other activities (e.g. leisure excursion and shopping long term) do not have enough samples in the data to appropriately fit the model.

These results show that the simulation is capable to reproduce most observations and that it produces appropriate variability by using the above model configuration. At the same time, it is necessary to examine the decision heuristic in more detail to clarify whether the execution of 23 hours of work is what the model defines or a flaw in the code. Moreover, additional simulation runs are necessary in order to sample from a larger dataset.

5 Outlook

The next task is to test our model on a population of same agent types (e.g. married working male between 30 and 40 years of age). Although we think that it is more difficult to reproduce the behavior of a specific agent, we still need to demonstrate that our model can reproduce observed behavior of agent types without overfitting the data sample. We also need to validate the location choice procedure of our model. The current approach might not be able to reproduce real location choice distributions observed over longer time intervals. Horni’s findings (Horni et al. (2011)) might provide helpful insights for this task. In combination with effectiveness functions, it could become a simple but effective location choice model that is also able to reproduce seasonal effects, something that is important for a continuous model.
Figure 6: Illustration of the statistics for activity duration, frequency and percentage of time spent per week extracted from the simulation runs (see Section 4.1 for an explanation how to interpret the box-plot).

(a) **Base case.** See Section 4.2 for a detailed explanation.

(b) **Exact case.** See Section 4.3 for a detailed explanation.

(c) **Use case.** See Section 4.4 for a detailed explanation.
Validation of a Continuous Simulation Model for Daily Travel

Figure 7: Illustration of the statistics for similarity, execution probability and transition probability extracted from the simulation runs (see Section 4.1 for an explanation how to interpret the tables).

(a) **Base case.** See Section 4.2 for a detailed explanation.

(b) **Exact case.** See Section 4.3 for a detailed explanation.

(c) **Use case.** See Section 4.4 for a detailed explanation.
Figure 8: Illustration of statistics expressing the scheduling quality of our model (see Section 4.1 for a detailed explanation). The plot shows the start/end time distribution of work and the box-plot the execution interval (waiting time between activity executions) of activities.

(a) **Base case.** See Section 4.2 for a detailed explanation.

(b) **Exact case.** See Section 4.3 for a detailed explanation.

(c) **Use case.** See Section 4.4 for a detailed explanation.
6 Conclusion

This paper validates a microscopic travel demand simulation that can continuously simulate agent’s behavior. The continuous nature of the simulation will enable an investigation of traffic effects that occur between days and between weeks. Behavioral targets are central for the model. These targets are closely related to statistical data provided by various sources (e.g. Swiss Federal Statistical Office (2006)), simplifying model utilization for practitioners. Time-dependent effectiveness functions model various effects like shop opening hours or social and cultural norms. Agents keep track of their performance and compare it to behavioral targets. Deviations cause discomfort which is conveyed to a planning heuristic, making on the fly decisions about upcoming activities agents should execute. Three model configurations illustrate different aspects of the model. The first configuration performs model calibration and shows that it reproduces underlying measures without additional constraints. The second configuration investigates how agents can be brought to execute specific activities. This ability is important for e.g. social interactions, requiring that agents comply with their appointments at specific places and times. The third configuration extracts a behavioral pattern from a six-week continuous travel diary, representing a person’s weekly rhythm. The validation of the simulation using several statistics and measures demonstrates that the model reproduces various behavioral aspects observed in the data.

7 References


