Simulating the influence of Social Contacts Spatial Distribution on Mobility Behavior

Thibaut Dubernet
Kay W. Axhausen

April 2015

STRC
15th Swiss Transport Research Conference
Monte Verità / Ascona, April 15 – 17, 2015
Simulating the influence of Social Contacts Spatial Distribution on Mobility Behavior

Thibaut Dubernet  
Institute for Transport Planning and Systems (IVT)  
ETH Zurich  
CH-8093 Zurich  
phone: +41-44-633 68 65  
fax: +41-44-633 10 57  
thibaut.dubernet@ivt.baug.ethz.ch

Kay W. Axhausen  
Institute for Transport Planning and Systems (IVT)  
ETH Zurich  
CH-8093 Zurich  
phone: +41-44-633 39 43  
fax: +41-44-633 10 57  
axhausen@ivt.baug.ethz.ch

April 2015

Abstract

In their daily life, individuals are frequently involved in joint decision making — situations where several individuals have to agree on the actions they will undertake to achieve a joint outcome. Examples in the context of mobility behavior include intra-household task allocation, intra-household vehicle allocation, choice of the time and venue for a dinner with friends or traveling together in the same private vehicle.

In addition to being necessary to predict joint travel and car occupancy, it has been hypothesized that considering explicitly this kind of joint decision process for the case of leisure activities planning might help to improve the forecasts for the choice of the leisure destination, due to the often social nature of such activities. This research is motivated by those results, and aims at including social behavior — including joint decision making and coordination — in a multi-agent transport simulation in a meaningful way. To do so, an algorithm to solve a specific game-theoretic solution concept has been designed.

This framework is used to simulate joint leisure activity location choice. A synthetic social network, generated using the approach of Arentze et al. (2013), is used to obtain realistic geographical distribution of social contacts. Validity of the predicted traveled distances, as well as sensitivity of the results to the geographical properties of the social network will be demonstrated.
1 Introduction

In developed countries, a continuous increase of the share of trips which are performed for leisure purposes could be observed in the last dozens of years (Schlich et al., 2004; Axhausen, 2005). This represents a challenge for travel behavior modeling, as those trips are much more difficult to capture than commuting trips: they are performed more sporadically, and data about those trips is much more difficult to collect. Understanding better how destination choice for leisure trips is made is therefore essential to improve the accuracy of those forecasts. This increase in leisure travel has been anticipated early, and the social nature of such travel already hypothesized, for instance by Salomon (1985), who stated that “one particular type of travel, that for recreational and social purpose, may increase when more leisure time is available”. This pronostic was later confirmed, for instance by Stauffacher et al. (2005), who analyzed the motives behind leisure activities, using the results of a 12 weeks leisure travel diary survey. They found social contact to be the most important, and that in addition respondents travelled with social contacts for more than 70% of leisure activities. This fact, between others, generated a growing interest in the social dimension of travel, and how travel decisions are influenced not only by the global state of the transportation system, but also by joint decisions and interactions with social contacts — a clear sign for this interest being the regular workshops organized on this theme (Dugundji et al., 2008, 2011, 2012).

Previous studies have been conducted with the idea that an important factor in leisure trip destination choice, or activity duration choice, is the ability to meet social contacts. Examples of empirical work include Carrasco and Habib (2009), Habib and Carrasco (2011) or Moore et al. (2013). All those studies show a significant influence of social contacts on the spatial and temporal distribution of activities. In addition, the influence of the social nature of human beings was shown to generate paradoxical effects. For instance, Harvey and Taylor (2000) show that persons working from home tend to travel further for leisure purpose, in order to fulfill their need for social contact, that they cannot fulfill at their workplace. A model ignoring such effects might thus substantially underestimate the traveled distances for such individuals.

A very active field of research is the study and modeling of intrahousehold interactions and joint decision making, often using the classical random utility framework extended to group decision making. A classical way to cope with the possibly conflicting objectives of different members of the household is to specify a group level utility function. For instance, Zhang et al. (2005, 2007) develop a model where time for different activity types is allocated to household members, subject to time constraints (including equality of time participation in joint activities), using a group level utility function formulated as a multilinear combination of the individuals’ utilities — that is, a linear combination of individual utilities and pair-wise product of individual utilities.
Kato and Matsumoto (2009) use a linear combination of the utility functions of the household members as a group utility. The assumption behind this kind of models is the existence of “utility transfers”: individuals accept to decrease their own utility if it allows to increase the utility of others by a certain fraction of their loss. Bradley and Vovsha (2005) focus on the “daily activity pattern” generation, with household “maintenance” tasks (e.g. shopping) allocation and possibility of joint activities. To do so, they assume a layered choice structure, choosing first a daily activity pattern for each member, and then assigning joint and maintenance activities. Gliebe and Koppelman (2005) also base their model on the daily activity pattern concept, choosing first a “joint outcome” (the sequence of individual and joint activities), and then an individual pattern for each household member. Those models rely on enumeration of the possible household level patterns. Gliebe and Koppelman (2002) also derived a constrained time allocation model, which predicts the time passed by two individuals in joint activities. Rather than postulating a group level utility function, the models of those authors specify a special distribution for the error terms of the individuals. In this setting, the error term of the individuals are correlated so that the probability of choosing a given joint output is the same for all individuals. Ho and Mulley (2013) also estimate models in which members of the household perform choices constrained by the choice of a household level travel pattern. Their data, as well as the parameters of the models, show high joint household activity participation on weekends, and a high dependence of joint travel on trip purpose and household mobility resources. Those results highlight the importance of representing joint household decisions, in particular when going beyond the “typical working day”. Vovsha and Gupta (2013) formulate a time allocation model for multiple worker households, which considers a positive utility for members of the household to be home jointly, as it makes joint activities possible. The estimation results show a significant influence of this kind of synchronization mechanism. Most models listed in this paragraph are specific to given household structures; in particular, separate models need to be estimated for different household sizes.

Household level decision processes have also been modeled with approaches which significantly differ from the classical random utility framework. Golob and McNally (1997) propose a structural equation model, which predicts time allocation and trip chaining based on the sociodemographics of a household. Golob (2000) also used a structural equation model to model the dependency of time allocations of the two heads (man and woman) of a household.

Another class of approaches, more oriented toward multiagent simulation than analysis, is the use of optimization algorithms to generate households plans. They handle the household scheduling problem by transforming it into a deterministic utility maximization problem. Contrary to the previously presented approaches, those alternatives do not lead to the estimation of a model against data. The first of those approaches was introduced by Recker (1995). By extending increasingly the formulation of the Pick-Up and Delivery Problem With Time Windows, a
well studied combinatorial optimization problem, he formulates the problem of optimizing
the activity sequence of members of a household as a mathematical programming problem.
However, due to the complexity of the problem, the full problem cannot be solved exactly by
standard operations research algorithms, and the activity durations are not part of the optimized
dimensions. Chow and Recker (2012) designed an inverse optimization method to calibrate
the parameters of this model using measured data. Also, the formulation from Recker (1995) was
later extended by Gan and Recker (2008) to introduce the effects of within day rescheduling due
to unexpected events. Another attempt to generate plans for households uses a genetic algorithm,
building on a previous genetic algorithm for individual plan generation (Charypar and Nagel;
2005; Meister et al., 2005a). This algorithm optimizes sequence, duration and activity choice for
a household, rewarding the fact that several members of the household perform the same activity
simultaneously, in the way also used by Yoksha and Gupta (2013). Finally, Liao et al. (2013)
formulate the problem of creating schedules for two persons traveling together as finding the
shortest path in a “supernetwork”, and solve this problem using exact shortest path algorithms.
They however note that their model is specific to the two person problem, and that extension
to larger numbers of agents may prove to be computationally expensive. All those approaches
remained experimental, and were not integrated into multiagent simulation tools.

Another class of methods aiming at multiagent simulations consists of rule based systems, which
use heuristic rules to construct household plans. Miller et al. (2005) develop such a model for
household mode choice. The main difference with an individual mode choice model is the
consideration of household level vehicle allocation. In their model, individuals first choose
modes individually. If a conflict occurs, the allocation that maximizes the household level utility
is chosen. The members which were not allocated a vehicle will fall back on their second best
choice, and/or examine shared rides options. Arentze and Timmermans (2009) develop a
rule base model which relies on a simulated bargaining process within the household. Though
such models can easily represent complex decision processes, their calibration and validation is
cumbersome.

Another field of empirical research studies the spatial characteristics of social networks. For
instance, Carrasco et al. (2008) studied the relationship between individual’s socioeconomic
characteristics and the spatial distribution of their social contacts. This kind of empirical
work allows to specify and estimate models able to generate synthetic social networks, given
sociodemographic attributes and home location. Another kind of data collection is the one of
Kowal (2013), that uses the technique of snowball sampling, where random individuals are
asked to list social contacts, that are in turn contacted and asked the same set of questions. Based
on this data, Arentze et al. (2012) estimated a model capable of synthesizing social networks
with realistic geographical and topological properties. This kind of model is essential if one
wants to include social network interactions in microsimulation model.
This integration of social networks in multiagent simulation frameworks has already been attempted by other authors. Due to their disaggregated description of the world, such models are particularly well suited to the representation of complex social topologies. Han et al. (2011) present experiments of using social networks to guide activity location choice set formation in the FEATHERS multiagent simulation framework. Using a simple scenario with 6 agents forming a clique, they consider the influence of various processes like information exchange and adaptation to the behavior of social contacts to increase the probability of an encounter. They do not, however, represent joint decisions, such as the scheduling of a joint activity. The same kind of processes have been investigated by Hackney (2009), using more complex network topologies, within the MATSim framework, used in this paper. Ronald et al. (2012) and Ma et al. (2011, 2012) present agent based systems which do integrate joint decision making mechanisms, based on rule based simulations of a bargaining processes. Frei and Axhausen (2011a) demonstrate a simple joint planning model, where (a) social contacts decide to perform a joint activity if it improves the utility of all co-participants (b) location of a joint activity is chosen to maximise a group utility. They are not yet integrated into any operational mobility simulation platform.

This data allowed the estimation of social network generation models, to create the input for microsimulation softwares. Illenberger et al. (2009) propose a simple model for generating synthetic social networks, based on the data collected for Zurich (Kowald and Axhausen, 2012; Kowald, 2013). Frei and Axhausen (2011b) compare two approaches for social network generation, on toy examples (with points randomly located in a square). Arentze et al. (2013) developed a more sophisticated social network generation model, that not only uses home location as an explanatory variable, but also gender and age similarity, and includes explicitly transitivity — the fact that, everything else equals, two individuals have a higher probability to be friends if they have common friends.

Meister et al. (2005b) did a first attempt to simulate joint planning in an household, using the concept of a joint household utility. Dubernet and Axhausen (2014) integrated this concept into the MATSim framework, showed that this simple idea produces unrealistic behavior, and proposed another solution concept for the joint planning problem, seen as a game theoretic game. This solution framework was shown to behave pretty well for the household case (Dubernet and Axhausen, forthcoming), and its application to more general social networks is being explored.

Those remarks point the need to represent social contacts in microsimulation, to actually represent the influence of social contacts, and of their geographical distribution, on travel behavior. This paper presents a model to represent joint decisions, that is a prerequisite for testing such effects. The variability of the results under different social network characteristics is then explored, for a scenario in the Zurich area.
2 Model and Simulation Framework

The current section presents a simulation framework for the simulation of joint decisions for mobility behavior forecasting.

This section gives a brief overview of the theoretical underpinnings of the simulation framework. More details can be found in Dubernet and Axhausen (2012, 2013b,a) or Dubernet and Axhausen (2014).

Game theory, as a theoretical framework to represent competition, has been used in many forms in transportation research. One of the earlier examples, and probably one of the most influential, is the Wardrop equilibrium condition in traffic assignment (Wardrop, 1952), which is simply a Nash equilibrium of a specific congestion game. This equilibrium notion has then widely spread in transportation research in general, and traffic assignment in particular, and doing an exhaustive review is not the purpose of this paper.

Although the outcome of any game is a decision "joint" in some way (the decision of a player depends on the decisions of the other players), this work uses a more restrictive definition of what is a joint decision.

A joint decision, as we understand it here, is a set of interlinked decisions by several players, requiring the usage of explicit coordination, or binding agreements. Including such possibility in a game theoretic framework requires a specific solution concept.

This can be illustrated by a classical game, called the House Allocation Problem (Schummer and Vohra, 2007). This game consists of \( n \) players and \( n \) houses. Moreover, each player has its individual ordering of the houses, from the most preferred to the least preferred, and players prefer being allocated alone to any house rather than in the same house as somebody else. The strategy of a player is the house chosen to live in.

An interesting feature of this game is that any one-to-one allocation of players to houses is a Nash Equilibrium: no player can improve its payoff by unilaterally changing its strategy, as it would require choosing an occupied house. This result however contradicts basic intuition about the stability of such an allocation. In this particular case, a more realistic solution concept is the Absence of Blocking Coalition: given a one-to-one allocation of houses to players, a blocking coalition is a set of players which could all be better off by re-allocating their houses among themselves.

It is to be noted that both solution concepts correspond to rational agents, i.e., agents having a
preference ordering over outcomes. What differentiates both solution concepts is the degree of communication which is hypothesized: in a Nash Equilibrium, for a given player, the strategies of the others are taken as given; in an Absence of Blocking Coalition, players have the possibility to "negotiate" a change of strategy with other players, which will be accepted only if all agents in the negotiation are better off after the re-allocation. In this work, we consider that a Nash Equilibrium corresponds to individual decisions only, whereas the blocking coalition concept allows what we name joint decisions.

Given those remarks, a solution concept for the "daily planning game", including the possibility of binding agreements, is adopted: the Absence of Improving Coalition concept. Given an allocation of daily plans to individuals, an improving coalition is a set of social contacts that could all be better off by simultaneously changing their daily plan. One can think of a group of friends switching from individual dinners at home to a joint dinner in a restaurant.

A co-evolutionary algorithm is designed to solve this problem, using the MATSim framework. The basic modeling idea is that individuals associate a utility value to their day, which increases with the time spent performing activities and decreases with the time spent traveling. Different parameters can be used for different modes or activity types, using the functional form from Charypar and Nagel (2005). Travel time is influenced by other agents via congestion. Co-evolutionary algorithms are particularly well suited for this kind of problems (Ficici, 2004; Popovici et al., 2012).

Figure 1: The MATSim iterative process

![Figure 1: The MATSim iterative process](image)

The co-evolutionary algorithm used to solve this problem is an emulation of a learning process (Nagel and Marchal, 2006). Using the iterative learning analogy, the specification of the algorithm is quite natural: each agent will perform an evolutionary algorithm to optimize its own daily plan, the fitness of which will be evaluated by executing all daily plans on the network to evaluate the resulting state of the transportation system. The steps of this process are represented on Fig. 1.

The first step is the specification of an Initial demand. All agents have an initial daily plan, which will serve as a starting point for the iterative improvement process. Some characteristics of the
plans are left untouched during the simulation, and should therefore come from data or an external model. This is typically the case of long term decisions, such as home and work locations, or decisions involving a larger time frame than a single day (e.g. do the weekly shopping or not).

The second step is the Mobility simulation. Plans of all agents are executed concurrently, to allow estimating the influence of the plans of the agents on each other. This step typically uses a queue simulation to simulate car traffic, which gives estimates of the congested travel time. Simulation of bus delays due to congestion and bus bunching can also be included. Together with the next step, this step constitutes the evaluation stage of the co-evolutionary algorithm.

Then comes Scoring. The information from the simulation is used to estimate the score of each individual plan. This information typically takes the form of travel times and time spent performing activities; experiments also included information such as facility crowding (Horni et al., 2009). The functional form is the one used by Charypar and Nagel (2005). It uses a linear disutility of travel time, and a logarithmic utility of time passed performing activities. Different parameters can be defined for each mode/activity type.

This gives the score from a single interaction. The fitness of the daily plan (entity of the algorithm) can then be updated, as \((1 - \alpha) f_{\text{old}} + \alpha f_{\text{new}}\), with \(\alpha \in [0.5, 1]\) being the learning rate. The lowest the learning rate, the more the fitness of a plan will be close to an average fitness over the evaluated interactions. While this is consistent with the hypothesis that individuals react to the expected state of the transport system, most applications use a learning rate of 1, which results in more reactive agents, and thus faster convergence.

The last step is Replanning. This step actually groups two of the important components of co-evolutionary algorithms: (a) selection of the interactions for evaluation, and (b) application of the evolutionary operators (selection and mutation). To do so, part of the agents select a past plan based on the experienced score, following a Logit selection probability. This will have two consequences: (a) the state of the transport system, used for evaluation, will only evolve slowly from iteration to iteration, giving the time to the agents to adapt, and (b) those plans will be re-evaluated, given the new plans of the other agents. The other agents copy and mutate one of their past plans. What kind of mutation is performed determines which alternative plans will be tried out by the agent.

Those steps are iterated until a stationary state is reached, and the state of the system in this stationary state is taken as a result.

Given this general framework, to be able to implement an algorithm searching for states without blocking coalitions, one needs a way to represent the influence of explicit coordination on the utility of a daily plan. This is solved by including joint plans constraints. A joint plan is a set
of individual plans executed simultaneously. Different copies of the same individual plan can be part of different joint plans — for instance an agent might go to a given restaurant alone, with members of its household or with a group of friends. The score of the different copies will take into account the influence of the joint plan to which it pertains. Those joint plan constraints are included using heuristic rules, applied after mutation operators are applied, and are classified as strong or weak constraints — weak constraints are considered when selecting plans for execution, but are allowed to be broken when merely selecting plans for mutation. They are then part of the evolution process. In the current application, the heuristic rules consist in joining newly created plans with joint trips (strong) or with leisure activities at the same location at the same time (weak).

To allow handling joint plans, replanning needs to be performed for groups of agents: agents are handled with all agents with whom they have a joint plan, plus some social contacts with whom new joint plans can be created, chosen randomly among the social contacts.

For each group, two actions are then possible. For most groups, an allocation of existing plans, fulfilling the joint plans constraints, is selected for execution. Based on plan scores, randomized by adding an extreme value distributed error term, an allocation without improving coalitions is searched for by an algorithm inspired by the “Top Trading Cycle” algorithm used for the House Allocation Problem (Schummer and Vohra, 2007).

For the other groups, a plan allocation is selected and copied. The copied plans then undertake mutation, to make the agents explore new alternative joint plans. What kind of mutation is performed determines which alternative plans will be tried out by the agent. The modules used in this study are:

- Departure time mutation
- Subtour mode mutation and re-routing
- Joint trip insertion/removal and re-routing
- Swap two random activities and re-routing
- Choose new leisure location for a group of social contacts

Agents have a limited memory size, keeping at most 3 plans per joint plan composition, and 10 plans in total. If this limit is exceeded, one should keep the plans which have the highest probability to create improving coalitions, that is, to be preferred to the other plans in the agent’s memory. To this end, a lexicographic ordering is used: the process removes the joint plan which maximizes the number of individual plans which are the worst of the agents’ memories. If several joint plans have the same number of worst plans, the process chooses among them the joint plan which maximizes the number of second worst plans, and so on until the "worst"
joint plan is unique. When the overall maximum number of plans in the memory of an agent is reached, the worst individual plan for this agent is removed along with plans of other agents of the same joint plan. Each agent keeps at least one plan not part of a joint plan, as there may otherwise not be any state without blocking coalitions. Agents are parsed in random order, to avoid the emergence of "dictators" over iterations, whose worst plan would always be removed, even if it is the only "bad" plan of a joint plan.

Though those selection operators seem to be in accordance with the chosen solution concept, it is difficult, if not impossible, to prove that the process will actually converge towards the state searched. As noted by Ficici et al. (2005), when they perform a theoretical analysis of different selection methods in a co-evolutionary context, "Co-evolutionary dynamics are notoriously complex. To focus our attention on selection dynamics, we will use a simple evolutionary game-theoretic framework to eliminate confounding factors such as those related to genetic variation, noisy evaluation, and finite population size". Those "confounding factors" can however not be eliminated from an actual implementation of a co-evolutionary algorithm, and rigorously proving that a given algorithm actually implements a specific solution concept is very tedious, if not impossible.

With iterations, agents build a choice set of daily plans that becomes better and better given the actions of the other agents. However, the presence of a large portion of agents with plans resulting from random mutation creates noise, not only for the analyst looking at the output of the simulation, but for the agents themselves when they compute the score of their plans. To solve this issue, when the system reaches a stable state, agents stop performing mutation, and only select plans from their memory for 100 iterations, using the absence of improving coalition with randomized scores. This ensures that the selected plans are the result of a behavioral model, rather than the result of random mutation operators.

3 Results

This section presents results for runs with two distinct social networks, for simulations in the Zurich Area.

Table 1 presents statistics of the synthetic social networks, compared with the statistics from the snowball sample from Köwäld (2013). The social networks were generated using the model from Arentze et al. (2013), estimated on this same dataset. This model works by testing the probability, for each agent in turn, to establish friendship with each member of a sample of the full population. The two social networks were generated using two different sampling rates:
0.025% for the long ties, and 10% for the short. Thus, homophily (the proportion of ties linking agents falling within the same age category or gender) is also lower for the social network with long ties. The social network is then restricted to the persons part of the Zurich-centered initial demand. This results in a decrease of the ego-alter distances, with only a very minimal decrease in degree.

Table 1: Characteristics of the Synthetic Social Networks

<table>
<thead>
<tr>
<th>Social Network</th>
<th>Clustering</th>
<th>Avg. Degree</th>
<th>Age</th>
<th>Gender</th>
<th>Distance (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snowball</td>
<td>0.206</td>
<td>22</td>
<td>46.3</td>
<td>61.7</td>
<td>26.6</td>
</tr>
<tr>
<td>Long</td>
<td>0.190</td>
<td>22</td>
<td>30.7</td>
<td>56.5</td>
<td>49.1</td>
</tr>
<tr>
<td>Long (ZH)</td>
<td>0.187</td>
<td>20.6</td>
<td>29.4</td>
<td>55.7</td>
<td>17.8</td>
</tr>
<tr>
<td>Short</td>
<td>0.150</td>
<td>21.7</td>
<td>45.2</td>
<td>66.0</td>
<td>18.8</td>
</tr>
<tr>
<td>Short (ZH)</td>
<td>0.225</td>
<td>20.6</td>
<td>45.4</td>
<td>66.2</td>
<td>7.3</td>
</tr>
</tbody>
</table>

Fig. 2 presents the traveled distribution per mode. The realism of the social network improves a lot the realism of the traveled distances for joint travel, in particular for the “driver” mode: drivers perform much less detours when social contacts are properly located. In addition, it makes the traveled distances for bike and walk shorter, improving their fit of the observed data, by adding more joint trip opportunities for trips too short for public transport, but too long for walk or bike.

Fig. 3 shows the distance distribution per Origin/Destination activity type pair. The geography of the social network has here only a minimal influence on the travelled distribution. This improvement in the prediction of the traveled distance was one of the motivations of this work. The absence of strong effect does not however means that the approach is not suitable: as for the low share of joint trips, this might come from the absence of correlation when sampling plans for social contacts, leading to too few joint leisure opportunities.

4 Conclusion

The geography of social contacts is assumed to be an important factor influencing daily mobility, as previous studies showed that leisure activities are mainly performed for social purposes. However, current forecasting tools fail to represent this kind of phenomenon.
Figure 2: Travel Distance Distribution per Mode

Figure 3: Travel Distance Distribution per Purpose
Representing the influence of the spatial distribution of social contacts on the characteristics of travel requires representing how individuals agree on a joint outcome.

This paper presented an algorithm to simulate this kind of decisions, based on an equilibrium formulation allowing coordination, that is solved through a *co-evolutionnary algorithm*.

Testing the algorithm with two social networks, with realistic and too long ego-alter home-home distances. There is an important effect of the realism of the social network on the realism of traveled distances as a car passenger. However, the willingness to perform joint activities does not seem to improve the travelled distances to leisure the way it was expected.

It is hypothesized that the remaining problems come from an inaccurate *population synthesis*, that includes also the allocation of activity chains to agents. Solving this problem is the most important of the next steps. A new survey, as well as the use of phone call data, is envisionned to develop a method to co-generate activity chains for social contacts.

5 References


