Working Paper

Collective Location Choice Model

Author(s):
Frei, Andreas; Axhausen, Kay W.

Publication Date:
2011-03

Permanent Link:
https://doi.org/10.3929/ethz-a-006562059

Rights / License:
In Copyright - Non-Commercial Use Permitted
Collective Location Choice Model

A. Frei and K.W. Axhausen
Abstract

In this paper, a methodology for incorporating social networks into models leisure activity location choice is discussed. Empirical work is discussed in order to establish a foundation for the methodology, and a hierarchical approach is presented. A simulated spatial social network based on empirical data is used for a collective location choice model. The results suggest that using a joint location choice model based on a social network can explain to some extent the gap between observed leisure trip length and the distance of the nearest leisure activity location measured in micro simulations.

Keywords

Social network, distance, geography, spatial, location choice, collective behaviour

Preferred citation style

1 Introduction

This research explores collective decision-making in social networks applied to a model of activity location choice for leisure activities and the participation in joint activities. Social networks pertain to the decision of resource allocation in planning and executing leisure activities, the manifestation of such decisions in the travel patterns of daily lives and the less frequent journeys outside the daily activity space. Social network structures have an important influence on a realistic representation of travel behavior.

Transport planning aims to describe, understand and model the choices people make during the execution of their daily lives, including the more or less frequent journeys outside their daily activity space (de Dios Ortúzar and Willumsen [2001]).

The dominant paradigm for transport planning work is the individual satisfying his or her needs while maximizing the utility derived from the activities undertaken. Disaggregated models are now available for practical application. As part of these developments, new problems arise and it has become clear that the location choice for executing these activities cannot be understood with just the attributes of the location, the available infrastructures to access the locations and the traveler’s characteristics without understanding how social networks of relatives and friends are distributed and involved in the joint location choice process.

Activities and travel involving multiple persons, particularly in leisure travel, result from a collective decision process that requires its participants to find a way to allocate and arrange their individual needs along with those of the other person(s). On an abstract level, this means finding an appropriate modeling structure to represent the utility derived by each individual from participating in leisure activities and the collective effects of these individual utilities on joint decisions, such as the participation in a sports training with your favorite training partner(s) further away or participating in a training with less preferred training partner(s) at a closer location.

At a more basic level, this research involves explaining travel behavior patterns which are only understandable when collective decision making is used as a base process. For example, the leisure trip distance distribution in Switzerland has a mean of 11.1 km (Bundesamt für Raumordnung und Bundesamt für Statistik [2007]). The nearest leisure activity location of each leisure category has a mean of 1.1 km in a spatially detailed agent based micro simulation of the Swiss overall work day (Horni et al., [2009]). Even though the categories of leisure activities are aggregated and do not consider the differences of those locations, the difference between the nearest leisure activity location and the travel patterns which evolve from the actual activity location cannot be
explained by considering only the choice of each single person independent from one another. The consideration of a collective location choice between individuals participating in an activity together and considering the distance distribution of the social contacts (Within 120 km the distance distribution between egos and their alters has a mean of 16.5 km (Frei et al., 2009) could shed light on this difference.

2 Research objectives

This research aims to model the process which leads to a joint location choice for leisure activities. To do so three different steps have to be accomplished - a spatial social network structure, a process to describe the organization within this network of smaller activity-groups and a joint location choice model.

The first research objective is accomplished by generating a static geographical embedded social network. To do so, an exponential random graph model is used to reproduce the edge distance distribution observed in an empirical study in Zurich, Switzerland. For further details of the method used to generate the social network, see Frei and Axhausen (2011).

The second research objective is the social organization within the social network. A joint activity participation model will be used to model the participation of each individual in a certain activity with others and to build an activity group. A randomly selected group of individual can find or invent a focus around which to combine various others with whom they are tied in an activity. The focus gets smaller (larger), with decreasing (increasing) number of participants (or individuals in the social network of interest). The foci are substitutes for each other.

The third research objective is to model the collective or joint location choice. To choose a location to perform a specific activity, a utility aggregation function will be introduced (preference function). The utility for location \( a \), \( U_a = f(\alpha_i, U_i) \), is an aggregated function over all \( i \) participants in an activity. \( \alpha_i \) represents a weight on the individuals’ utility \( U_i \), which are among others’ network attributes.

3 Literature review

Even though the social context of traveling is rather under-researched, the importance of joint activity participation is evident and has been studied in the past.
Kostyniuk and Kitamura (1983) analyzed time use data and found that joint activities tend to have a longer duration than other similar (non-work) activities. Furthermore people participating in joint activities travel further to perform an activity, which would be consistent with the finding of Horni et al. (2009) (see page 3). Many researchers studied especially the effect of household attributes on joint activity travel. For example Jones et al. (1983) and Kostyniuk and Kitamura (1983) found that adults are strongly affected by the presence of children. Couples with children perform most joint activities at home, whereas couples without children are more likely to perform joint out-of-home activities. The employment status of both parts of couples influences the starting point of joint activities; couples where both are employed tend to choose a start of the location out of home and vice versa. The same researcher also found that the availability of a car influences positively more individual time-use patterns of couples. Fujii et al. (1999) found that people rated the time spent involved in joint activities higher than activities spent alone (non-work activities). Their time spend in joint activities was especially more “satisfying” and they chose, if possible, to allocate time to rather joint than independent activities. Van den Berg et al. (2010) used a social interaction diary to study the factors influencing the planning of social activities. They found that social activities scheduled later in the day are less likely to be routine. In contrast, social activities of longer duration and taking place in the weekend are more likely to be routine or preplanned. Harvey and Taylor (2000) studied the influence of the work location on joint activities with time use data. They argue, that people which work at home spend more time alone and therefore show a tendency to travel more to fulfill their needs to social interaction. Carrasco and Miller (2006) describe the joint activity participation with egocentric social structure effects (degree of a person), the use of communication technology and socio-demographic variables. They found, that people with a high egocentric social network degree are more likely to perform joint activities. Also the availability of communication technology such as telephone and the Internet lowers the cost of coordination and influences the participation in joint activities.

Srinivasan and Bhat (2008) also analyze time use data to explore joint activities. They find that joint activities are of longer durations, significantly likely to take place at the residence of other people, and often confined to certain time periods of the weekday. In addition, important differences in these characteristics are also observed. The desire to participate in activities with non-household members generates additional travel to pick-up and drop-off the activity companions. Finally, they argue that our understanding of joint activity participation and their manifestation in travel patterns is underdeveloped because of a lack of data.

To be able to compare the model of a joint activity participation model, empirical data would have to be needed, which shows at a minimum the group size distribution.
Further, the composition of these groups would be interesting, e.g. how many of the participants are friends with each other. Such data is to the knowledge of the authors not yet available. But at least the participant numbers should be higher than the number of accompanying travelers in Figure 1.

Figure 1: Number of accompanying travelers

Source: Schlich and Axhausen (2003)

Modeling of joint decision structures has been done in the past mostly by interactions between household heads. Fujii et al. (1999) modeled the allocation of time to in-home and out-of-home activities with other family members, with non-family members and alone using a production function paradigm. Golob and McNally (1997), Simma and Axhausen (2001) and Carrasco and Miller (2006) modeled interactions in a structural equation format, simultaneously accounting for the allocation of time to activities. Zhang et al. (2002) expressed household utility as an aggregation of the linear utilities of individual household members, weighted by the gender of the decision maker, using simultaneously unrelated regression (SUR) to derive the optimal level of time allocation to in-home, out-of-home individual and out-of-home shared activities. A similar and more extended approach is used by Gliebe and Koppelman (2000). Based on the idea of Townsend (1987) and Bhat and Koppelman (1993), they propose a utility-based model to explain the decision to allocate time to both joint and independent activity types. They offered an empirical model, structured as a two-level nested share model, in
which two household heads make an upper-level choice between joint maintenance, joint leisure or independent activity time allocation (see also Gliebe and Koppelman 2002). The independent activity is partitioned at the lower level into choices between at-home, subsistence, independent maintenance and independent leisure time allocation. The composite utilities of the independent activities of the individual decision makers are weighted using a parametrization based on their attributes, such as employment level.

Townsend (1987) forms the basis for the approach taken in this research. He developed a concept for the analysis activity patterns of households and their members. Household utility maximization is combined in an aggregation function of each individual’s preferences. In his empirical research he split the utility in three components related to interpersonal and personal parts to trade-off in activity and travel patterns: Efficiency, companionship and power/altruism. Efficiency is linked to the specialization and the efficiency gained through it. Companionship is the need of people to engage in activities jointly, and power is illustrated through e.g. higher income and therefore a higher influence in decision making. But power is also illustrated as a form of altruism, where a person engaged in an activity with a less powerful individual by helping them or fulfill other altruistic tasks in the free time to gain esteem.

Gliebe and Koppelman (2000, 2002) extended the approach of Townsend (1987) with the concept of Bhat and Koppelman (1993) to generate household activities with three different types of activities: consumption, work, household (maintenance) and leisure (discretionary) activities. The total utility of a household is defined as the joint utility of the household members e.g. \( U_T = f(\alpha_i U_i + \ldots + \alpha_j U_j) \), where the \( \alpha_i \)'s are weights of each individual from \( i \) to \( j \). The utility of each household member has two components. The first component is a combination of

\[
V_i = \beta_0 i + \beta_1 C_i + \beta_2 S_{x_i} + \beta_3 S_{w_i} + \beta_4 S_{h_i} + \beta_5 S_{h_{ij}} + \beta_6 S_{h_{ji}},
\]

where \( C_i \) denotes the consumption and \( S_{x_i} \) denotes the satisfaction of each leisure, work and household. The joint components reflect the degree of positive utility each individual gains from joint activity participation and are modeled as taste variations across individuals. The second component, the altruism effect, describes the total utility together with component one as follows: \( U_i = V_i + \beta_7 V_{j \ldots k-i} \), where \( V_{j \ldots k-i} \) describes the utility of all others participating in the joint activity. The satisfaction terms are modeled as a function of time spent, e.g. \( S = nT_x - T_x^2 \) where \( S \) convex. and the total time is limited by \( T_0 \leq T_x + T_h + T_l \).

This research concentrates on leisure activities only, but the context of joint household activities will be extended to general joint leisure activities. The influence factor \( \alpha \) in the aggregation function will be extended and the influence which is described by Townsend (1987) and Gliebe and Koppelman (2000) as a power relative to income and gets replaced with the approach described by Carrasco and Miller (2006), where the structural
influence will be tested with different approaches. Different structural measurements are reviewed in the following paragraph.

3.1 Network Structural Centrality Measurements

Network structural centrality measurements are described e.g. in Wasserman and Faust (1994), Hanneman and Riddle (2005) or in Monge and Contractor (2003). The basic idea is to measure the relative importance of each individual in a social network. There are four different approaches which are mainly used to measure the centrality, and each different measurement has a different meaning.

**Degree Centrality** The degree centrality is the simplest measurement and describes the number of total direct neighbors of each individual relative to the total network size: \[ C_D(v) = \frac{\text{deg}(v)}{n-1} \]. The degree centrality is interpreted as the intermediate probability of an individual to receive something which is flowing through the network, e.g. the risk to catch a virus, or the probability to receive certain information.

**Closeness Centrality** Degree centrality measures do only account for the immediate neighbors. A critical example of the degree centrality approach could be, that an individual might have a high degree, but all those individuals he is tied to are isolated from the rest of the network and so he is not very central even though he has a high degree. There are several measurements for closeness centrality. The most common two measurements are the mean path distance and the eigenvector of geodesic distance. The mean path distance measures the mean shortest distance \( d_G \) in social space between individual \( v \) and all other members of network \( V \): \[ C_C(v) = \frac{\sum_{t \in V \setminus v} d_G(v, t)}{n-1} \]. But similar to the degree centrality, the mean path measurement can be somewhat misleading in larger networks, as e.g. an individual with a short path to many members of a disconnected part of the network and long paths to the rest of the networks, can have a similar measurement as e.g. an individual with a medium distance to all members. It can be argued that the second individual described is more central in the sense, that he can reach more other individuals with the same amount of effort. The eigenvector is an approach to find a measurement on the overall network structure with paying less attention to local structures (e.g. Google’s rank page uses some kind of eigenvector centrality). The eigenvalue score of \( v_i \) is \( x_i \), where \( A_{ij} \) is the adjacency matrix. The centrality score for
Collective Location Choice Model

March 2011

$v_i$ is proportional to the sum of the scores of all his neighbors: $x_i = \frac{1}{N} \sum_{j=1}^{N} A_{i,j} x_j$. $N$ is the total number of individuals in the network. There might be different eigenvalues $\lambda$ for which a solution to this equation exists. As a centrality measurement, we are only interested in the greatest eigenvalue.

**Betweenness Centrality**  The betweenness centrality measures positional advantages, or power of an individual $v$, such that the shortest path from an other individual to an other individual goes through individual $v$. The idea is that individuals who are between others, might be able to use this as some form of power.

The betweenness centrality is the measurement of how many of the shortest paths $\sigma$ between other individual of the network $V$ go through individual $v$: $C_B(v) = \sum_{s,t \neq v \in V, s \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} / ((n-1)(n-2))/2$.

### 4 Implementation

The implementation of a joint activity and location choice model is grouped in three hierarchically dependent models: A social network model $(a)$, a joint activity participation model $(b)$ and a joint location model $(c)$. The dependence is static and hierarchical in the sense of a fixed and timely order, where $(b)$ depends on $(a)$ and $(c)$ depends on $(a)$ and $(b)$: $(a) \rightarrow (b) \downarrow \cdots (c)$. The timely, hierarchical dependence is not dynamic and will therefore not be true for certain situations, where these three structures/decision situations have dynamic interactions among each other on different levels.

For example, a certain activity will be uniquely located on a short or mid-term horizon, and the activity participation is therefore a function of the location. On a longer horizon, the social network might be influenced by the location choice (e.g. when people decide to move). In general this model structure assumes that the social network is relative stable over time. The mean duration for social contacts is based with a survival model, which is modeled with hazard function, which describes the exponential failure density, of the Zurich data to be around 8 years (see Figure 2) and therefore influence of short term decisions should be neglectable. The dependence of the joint location model on the activity participation model, and not vice versa, is chosen because the pool of possible locations for a certain activity is normally much bigger than the pool for possible activities for a certain location. For example, the choice to go to a movie gives us a relatively
high flexibility in location choice (in bigger cities), whereas if we chose to go to a movie theater, the activity options are very limited. The argument here is that the model process should reflect the choice process while maintaining high flexibility.

Figure 2: Survival function of the empirical duration of social contacts

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2}
\caption{Survival function of the empirical duration of social contacts}
\end{figure}

4.1 Basic assumptions

The utility function for the joint activity participation model and the collective choice model is a modification of the utility function proposed by Gliebe and Koppelman (2002) and Townsend (1987). \( U_T = f(\alpha_i V_i + ... + \alpha_j V_j) \), where the \( \alpha \)'s are weights of each individual from \( i \) to \( j \). The utility of each person is \( V_i = \beta_{1i} S_{li} + \beta_{2i} S_{lij} \) for all \( H=1 \) \( - \beta_{3ij} S_{lij} \) for all \( H=0 \), where \( S_{li} \) denotes the satisfaction of each leisure activity and \( H = 1 \) means that person \( i \) is connected to person \( j \) and vice versa, if \( H = 0 \). The joint component \( S_{lij} \) reflects the degree of positive utility each individual gains/loses from joint activity participation.
4.2 Joint Activity Participation Organization Model

The static social network generated by the model described in Frei and Axhausen (2011) is used to organize the population in activities.

The idea of the joint activity participation model is that a joint activity $A_i$ is more satisfying than an activity performed alone. A joint activity is defined as $A_i = \{n_i, ..., n_j\}$, whereas as the $\sum H \in A_i \geq 1$. Each $n_k \in A_i$ with $H = 0$ has a negative impact on the activity utility. This is chosen, because of the static social network, where it is assumed, that the probability of $n_i$ knows $n_k$ is high, but they are not maintaining a friendship.

The joint activity participation choice cannot be evaluated directly, because the simultaneously decision of each person affects the decision of the others. To get a stable result from the organizing process within the network to participate in a certain activity an iterative approach is chosen as follows:

1. 50% of the population are randomly picked to initiate activities.
2. Each person which has a direct connection to one of the persons who initiated an activity will obtain the information about this activity.
3. An initial choice for participating in one of the activities is made based on the physical distance to the person who initiated the activity.
4. Iterate: In each iteration 10% of the population is randomly chosen to make a new choice based on all of the available information, which includes the possibility a choice based on another person participating in an activity, who then changes the activity.

The iteration process will be continued till the overall utility stabilizes. The utility for each person is calculated by

$$V_i = \beta_0 + \beta_1 S_{li} + \beta_2 S_{lij, for all H=1} - \beta_3 S_{lij, for all H=0}.$$  

4.3 Collective Choice Model

The collective choice model is based on the generated social network and the activity participation model. Each leisure activity $A$ can be performed at each location $L = \{l_1, ..., l_i\}$, as there is no further categorization of leisure activities in this model. Each location $L$ is spatially fixed within $S$. The location choice for an activity $A$ is defined by the group utility $U_T = f(\alpha_i U_i + ... + \alpha_j U_j)$, where $U_i = V_i + \beta_4 V_{j=k-i}$ and $V_i = \beta_1 S_{li} + \beta_2 S_{lij, for all H=1} - \beta_3 S_{lij, for all H=0}$. The weights $\alpha$ are network structural.
measurements as described on page 8. The choice set of activity locations is created by including each in the activity participating persons three closest locations. This can account for the effects that not all information about locations is available for everyone, but that the groups information is based on its members.

5 Results

The stop criterion for the joint activity participation model was set to 2.5% relative change in the sum of the utility of the population $\sum V_i$. It took on average 273 iterations per run to reach the criterion. As there are no group size distributions for leisure activities known to the author, the resulting mean, median, std. deviation of the group size and mean distance between members of a group are calculated for different proportions of the parameters of the utility function. The utility function $V_i = \beta_{0i} + \beta_{1i}S_{i1} + \beta_{2i}S_{ij}, \text{for all } H=1 - \beta_{3i}S_{ij}, \text{for all } H=0$ consists of three parameters, where the intercept $\beta_{0i}$ is set to zero. The satisfaction terms $S_{i1,2i,3i}$ are set inverse proportional to the distance to the activity initiator and multiplied by ten. The parameter $\beta_{1i}$ is set to one. The parameters $\beta_{2i}$ and $\beta_{3i}$ are set so that the sum of each pair equals to one: $\beta_{2i} + \beta_{3i} = 1$. The results for different parameter proportions are presented in Table 1. It is not surprising, that the group size depends very strongly on the proportion of $\beta_{2i}$ and $\beta_{3i}$, as they affect the total utility each gets from the participation of others in the activity. But it also shows, that through the effect of a negative utility each one gets from an $H = 0$, the group sizes are substantially smaller than the mean degree of the underlying social network. For group sizes equal to one, the distance was set to 1.1, to make it comparable to the results in Table 2.

<table>
<thead>
<tr>
<th>$\beta_{2i}/\beta_{3i}$</th>
<th>Mean [-]</th>
<th>Median [-]</th>
<th>Std. dev. [-]</th>
<th>Mean distance within group members [km]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>8.6</td>
<td>5.6</td>
<td>3.4</td>
<td>19.8</td>
</tr>
<tr>
<td>3</td>
<td>6.6</td>
<td>5.2</td>
<td>3.0</td>
<td>16.6</td>
</tr>
<tr>
<td>2</td>
<td>5.8</td>
<td>5.1</td>
<td>2.2</td>
<td>13.4</td>
</tr>
<tr>
<td>1</td>
<td>4.2</td>
<td>3.8</td>
<td>0.5</td>
<td>11.8</td>
</tr>
<tr>
<td>0.5</td>
<td>2.3</td>
<td>2.6</td>
<td>0.2</td>
<td>8.2</td>
</tr>
<tr>
<td>0.33</td>
<td>2.1</td>
<td>2.4</td>
<td>0.2</td>
<td>8.0</td>
</tr>
</tbody>
</table>

For the choice model, a location set was randomly generated and placed uniformly within the square region of size $l \times l = 100 \text{km}$, so that the mean physical distance for each vertex to the closest location is 1.1 km. The group’s utility for the location choice is calculated as $U_T = f(\alpha_1V_i + \ldots + \alpha_jV_j)$, where the $\alpha$’s are normalized weights represented by the
network structure centrality measurements and the $V'_i$s are calculated in the same way as in the joint activity participation model. The resulting descriptive statistics for the proportion of $\beta_2i$ and $\beta_3i$, and the different $\alpha'$s used are available in Table [2]. The resulting distances to the chosen locations are close to the distances which are observed by Horni et al. (2009) for larger groups, but still are substantially larger than the mean distance to the closest locations for smaller more realistic mean group sizes. The different network structure measurements have an impact. The degree centrality measurement as the $\alpha$ results in shorter distances than the closeness centrality, which makes sense as a high degree centrality is a measurement for people which share people around themselves, whereas people with a high closeness centrality are people with rather fewer contacts but in between two groups, which makes the average distance larger. The betweenness centrality does not measure much of the spatial impact, as there are no information passed further than one degree.

Table 2: Descriptive statistics of the distance of each vertex to the chosen activity location in km

<table>
<thead>
<tr>
<th>$\beta_2i/\beta_3i$</th>
<th>Degree centrality Mean</th>
<th>Median</th>
<th>Closeness centrality Mean</th>
<th>Median</th>
<th>Betweenness centrality Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>10.4</td>
<td>6.2</td>
<td>12.6</td>
<td>6.3</td>
<td>11.1</td>
<td>6.1</td>
</tr>
<tr>
<td>3</td>
<td>8.6</td>
<td>5.8</td>
<td>9.8</td>
<td>5.9</td>
<td>9.0</td>
<td>5.8</td>
</tr>
<tr>
<td>2</td>
<td>7.0</td>
<td>5.4</td>
<td>7.6</td>
<td>5.5</td>
<td>7.2</td>
<td>5.3</td>
</tr>
<tr>
<td>1</td>
<td>5.6</td>
<td>5.3</td>
<td>6.4</td>
<td>5.3</td>
<td>6.1</td>
<td>5.2</td>
</tr>
<tr>
<td>0.5</td>
<td>4.6</td>
<td>4.3</td>
<td>5.1</td>
<td>4.4</td>
<td>4.9</td>
<td>4.3</td>
</tr>
<tr>
<td>0.33</td>
<td>4.4</td>
<td>4.3</td>
<td>4.6</td>
<td>4.3</td>
<td>4.5</td>
<td>4.2</td>
</tr>
</tbody>
</table>

CONCLUSION

This paper shows the impact of a joint location choice with a theoretical simulation. The gap between the distances of the average closest leisure location in Switzerland (1.1 km) and observed location choices (14.5 km) in the Swiss Microcensus can be explained to some extend with spatial social networks.

The magnitude of proportions of the different beta parameters are unknown and would need further survey work. Also the impact of other utility functions has not yet been tested. E.g. the satisfaction terms $S_{ij}$ could be modeled as a function of the number of persons involved, e.g. $S_{ij} = nS_{ij} - S_{ij}^2$ where $S_{ij}$ is convex. The $S_{ij}$ could also be modeled as taste variations across individuals.

Further there should be sociodemographical variables incorporated in future surveys to make such a model estimable.
References


Hanneman, R. A. and M. Riddle (2005) *Introduction to social network methods*, University of California, Riverside, CA.


