

# Spatial relationships and social Networks

## An iterative approach

**Conference Paper**

**Author(s):**

Kowald, Matthias; Frei, Andreas; Hackney, Jeremy K.

**Publication date:**

2008-09

**Permanent link:**

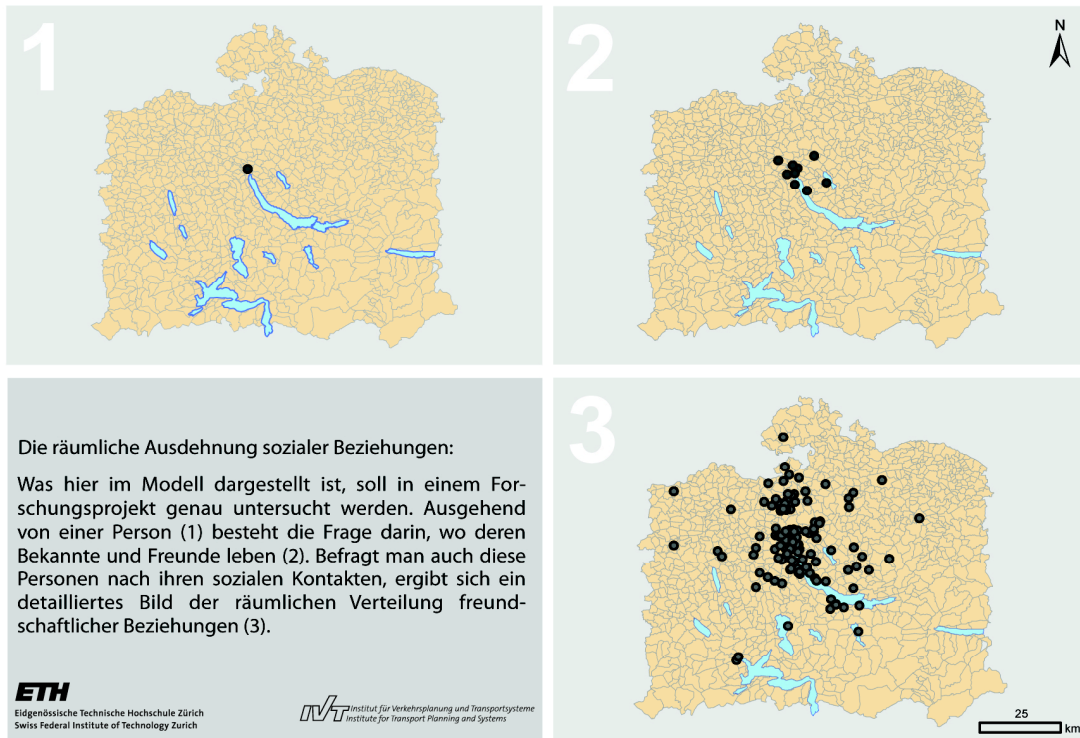
<https://doi.org/10.3929/ethz-a-005664674>

**Rights / license:**

[In Copyright - Non-Commercial Use Permitted](#)

**Originally published in:**

Arbeitsberichte Verkehrs- und Raumplanung 524



Conference Paper for “Applications of Social Network Analysis” (ASNA), Zurich, September 12-13, 2008

## Spatial Relationships and Social Networks: an Iterative Survey Approach

Matthias Kowald (IVT, ETH Zurich)  
 Andreas Frei (IVT, ETH Zurich)  
 Jeremy Hackney (IVT, ETH Zurich)

September 2008

Words: 6.700

Images: 5

**IVT** Institut für Verkehrsplanung und Transportsysteme  
 Institute for Transport Planning and Systems

**ETH**

Eidgenössische Technische Hochschule Zürich  
 Swiss Federal Institute of Technology Zürich

## **Spatial Relationships and Social Networks: an Iterative Survey Approach**

M Kowald

IVT

ETH

CH – 8093 Zürich

Telefon: +41 44 633 30 89

Telefax: +41 44 633 10 57

kowald@ivt.baug.ethz.ch

A Frei

IVT

ETH

CH – 8093 Zürich

Telefon: +41 44 633 41 02

Telefax: +41 44 633 10 57

frei@ivt.baug.ethz.ch

JK Hackney

IVT

ETH

CH – 8093 Zürich

Telefon: +41 1 633 33 25

Telefax: +41 44 633 10 57

hackney@ivt.baug.ethz.ch

September 2008

### **Abstract**

The planned survey aims to look behind social ego networks. Therefore a snowball sample will be taken. The technique of snowball sampling involves collecting data from ego networks, for example the names of ego`s social relationships, that are of interest for the investigation. But instead of only asking for the names of these contacts, the researchers also want to know their postal addresses. Knowing these addresses opens up the possibility of asking ego`s contacts if they also want to participate in the sample. So ego`s contacts become themselves ego`s, and so on. By using this method one not only collects single viewed ego networks, but mutually reported network structures. Because the network structure is surveyed from many points of view, the researchers will be able to see the indirect relationships between actors, which has not been possible with traditional ego network approaches that only capture direct relationships. Besides the spatial spread of social contacts, the survey will also collect data that deal with travel behaviour. For that reason there will be questions about how related actors stay in contact, when and where they meet and how they get to these places in terms of travel modes.

The survey will collect characteristics of ego and his contacts and a description of their relationships. The aims of the survey are to get experience with iterative snowball sampling techniques and to sample large geographically-embedded global networks. The collected data will be analysed and compared with data collected with traditional sampling methods.

### **Keywords**

Social network, snowball sample, ego-centered network,

### **Preferred citation style**

Kowald, M., Frei, A. and J.K. Hackney (2008) Spatial Relationships and Social Networks: an Iterative Survey Approach, *Paper for ASNA Conference Zurich 2008*, Zurich.

## **1. Abstract**

In a joint project, the Institute for Transport Planning and Systems (IVT) of ETH Zurich and the Institute for Sea- and Land-Transport (ILS) of TU Berlin plan to survey ego-centred social networks in an iterative process. Beyond this snowball survey method, the ego-centred network approach will be further developed by combining it with a personal diary for activities. The paper will give an overview of the planned survey and its' aims and provide a closer look with respect to the methods and strategies for data collection. Also, first empirical results from a pre-test will be presented.

## **2. The Link between social network analysis and transportation planning**

In the past, most work in transportation science used the single individual human being and its behaviour as the origin for explanations. This individual person is seen as an autonomous decision making actor. The picture is strongly influenced by the model of homo oeconomicus, which was developed very early in economics. Here every actor has the same attributes: He is perfectly informed about his needs and the opportunities in his environment. The actions of such a person are unilaterally oriented on the most rational way to satisfy its needs. To reach a full satisfaction level the individual has to move to different places. Travel phenomena emerge as an effect of the simultaneous movements of hundreds, thousands or millions of actors.

In recent years a new phenomenon made the development and use of new explanatory approaches necessary. It is the growing sector of leisure travel and the need to explain it. This kind of travel seems to be influenced not only by the needs of the individual human being but highly influenced by its social contacts. A better understanding of leisure travel is therefore connected with an understanding of how the social environment of a person is able to influence its decisions. „Accordingly, the distribution of these friends, relatives, acquaintances and contacts across space, or better space-time becomes crucial to an understanding of leisure travel and its potential for further growth“ (Axhausen, 2005, 90). Accordingly it becomes necessary to obtain data that imply information about single human beings and their social environment simultaneously.

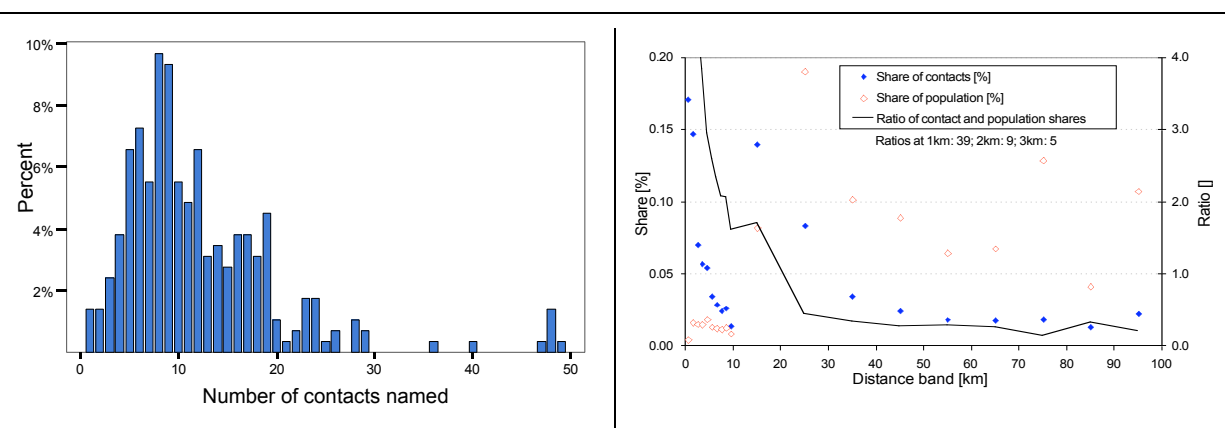
Further, the rising importance of the theory of social capital and further developments in this field steered transport planning towards usage of social network theory. A minimum definition describes social capital as resources that are not the property of an individual actor but can be generated in interactions between the members of a social network. The network of interest can be macroscopically or microscopically, the actors can be persons, institutions or states (see Jansen, 2003). To generate and obtain this resource the members of the specific network have to stay in contact. Only in this way it is guaranteed that there will be occasions for future exchanges to profit from the benefits of social capital. Using this idea as a basis there is a high likelihood that most persons in a specific social group do influence one another when making decisions about for example the living place, working place or where to spend holidays. To get an idea about how strong this influence is, the scientific field of transport planning is going to use the methods of social network analysis. Although explanations in transport planning are still using the concept of the maximisation of benefits, „The perfectly informed and unilaterally utility-maximising traveller is transformed into a network actor, who draws on the resources of his or her network but is constrained by its expectations while negotiating productive solutions for his or her daily life“ (Axhausen, 2008, 2).

In most studies where social network analysis is used, it is limited to an observation of social topology. In the case of its application in transport planning it will be combined with the geographical distribution of the social contacts of persons and the activities to see one another. This linkage is yet not well explored by sociologists or geographers. „Moreover, although data collection and modelling techniques have gone far toward understanding individual activity-travel decision-making processes in time and space, there is little known about the linkages between social and spatial interactions“ (Carrasco, Hogan, Wellman, Miller, 2007, 1). The usage of social network analysis will allow a detailed look on this connection.

### 3. Previous Work and new questions

In previous work, Frei and Axhausen explored ego-centred networks and mobility biographies for a random sample in Zurich City (see Frei and Axhausen, 2007). They found indicators for mobility biographies, by asking the respondents to report their previous living and working places. To explore the social network of the respondents they asked for a list of persons that have some kind of close relations with them. Whilst respondents were allowed to mention as many contacts as they wanted, there were on average 12.35 contact persons per respondent with a range from 1 to 49. Further questions showed that the share of core contacts decreases the higher the number of mentioned shared contacts is. Also Frei and Axhausen found that nearly two-thirds of the contact persons' home locations were located within 25 km from the respondents' home locations.

Figure 1 Ego-centred networks in Zurich City



Source: Frei and Axhausen (2007, 11 and 19)

There is a correlation between the living place of a person and the living places of its friends. Whilst this findings are verified by other papers (see Wellman, 1996) there is much need for further research to get a more detailed view on personal network structures. In a new research project there will be a closer look at the following aspects:

- What is the precise structure of personal networks?
- What kind of attributes do the persons of a network share, which do they not share?
- By which modes do persons stay in contact?
- To which proportion do a person and his contacts engage in joint activities?
- How and by whom are social interactions planned?

The new research project called “Travel impacts of social networks and networking tools” will combine an iterative ego-centred network approach with an activity based travel diary, to get information about these questions.

## **4. Asking a person about its network and activities**

### **4.1 The ego-centered network approach**

The first step in every social network analysis is to define a boundary for the network of interest. This can generally be done in two ways (see Jansen, 2003). The first is to collect information for a complete social network that is formally bounded by some kind of an administrative context. Examples for such units are companies, social institutions and districts in towns. This method is not appropriated for the questions guiding the interests of transport planning in the present study. In this project not only social relations within a specific administrative context are of interest but all kind of relations of a person influencing its behaviour. To explore the structure of these networks one can use the ego-centred network approach. The focus of interest lies on a person, called Ego, that reports from its point of view about the relations to other actors, so called Alters. He also gives information about the attributes of these Alters and about the way the different Alters are related to one another. So it is possible to investigate large personal networks. That would not be possible by using the method of exploring complete social networks as personal networks are not formally bounded.

#### **4.1.1 The name generator technique**

The network of interest is bounded by Ego's reporting who belongs to the network and who does not. To get the names of these Alters a questioning technique called name generator is used. The use of such a name generator has two aims. It helps to focus Ego towards the contacts that are of interest and it helps Ego to remember these contacts at all. It is a well documented phenomenon that the process of remembering concrete facts within the human brain works much better when supported by given stimuli. For that reason an unsupported collection of contact names is not recommendable (see Pool and Kochen, 1978).

A name generator exists of one or more questions towards Ego's social contacts differentiated by stimulus. The technique can be classified by the stimuli used. In the early years of its use, name generators often asked for names in a certain context. Often questions aimed for family-, work- or friendship-related alters. The main problem with questions of that type is that the context can be interpreted in many ways. For example the concept of friendship is not interpreted in a uniform way between cultures. Who can be seen as a workmate and who as a friend may even differ within one culture by social milieu. Because of this problem most of today's name generators deal with concrete interaction stimuli and not with contexts. Examples for questions are the frequency of visits in a given time frame or which person ego trusts when talking about fears or themes of highly importance.

Two often-compared stimulus based name generators are the instruments of Claude Fischer and Ronald Burt. In the early 1980s Fischer developed a name generator that is composed of eleven questions, each of them dealing with a special interaction stimulus (Fischer, 1982). The number of possible Alters elicited by the questions is unlimited. On the other hand, Burt developed a one-question name generator, dealing with the stimulus of discussing important issues (Burt, 1984). Here the number of possible Alters is limited to five. However, which generator is used has great implications for the collected data. From a quantitative point of view it can be said that on average the bigger the generator used, in the sense of more questions and therefore stimuli, the bigger the network under investigation will be. While on average respondents report 18.5 contacts when answering the questions of the Fischer instrument, there are only 3.0 contacts in the case of Burts instrument. "A single name-generator causes small ego-centric social networks. The size of the networks surveyed with the Fisher-instrument are remarkably higher then those with the two other instruments" (Frei and Axhausen, 2007, 11). These results make clear that the reported network size is correlated with the name generator used. But there are also qualitative data implications within the name

generator concept. At first it is possible to compute the multiplex-coefficient when giving more than one stimulus. A multiple relation means that a certain alter is an important person with regard to more than one stimulus for Ego. Another advantage of the multi-stimulus approach lies in the chance to remember and mention a forgotten Alter in the questioning process. A forgotten Alter will be missing in the study when using only one question. The Fischer approach provides the chance to remember and mention this person to an additional stimulus. In the case of using a multi-generator only an indication for multiplexity will be missed, but not the person. Comparing the results from measurements of a known, complete network and one with some missing actors, it becomes clear “(...) that using network level measures based only on alters elicited by a single name generator results in measurement error. Results from using network level measures to predict the probability of alters being recalled might further suggest that the size of this measurement error in estimates of network size would not be random, but related to properties of the network“ (Marin, 2004, 303).

Besides the higher number of mentioned contacts and the higher qualitative level of the surveyed data, there is also a disadvantage that goes hand in hand with the use of a multiple name generator. It is the burden that arises from the questions for the respondent. The more questions this person has to answer, the more time it will take and the higher is therefore the burden.

New approaches try to combine the advantages of the Fischer instrument with the lower response burden of the Burt instrument. One of these instruments, a so called modified multiple name generator, was developed and tested at the Institute for transport planning and systems (IVT) at ETH Zurich in 2006 (Frei and Axhausen, 2007). Here, four prompts are used in three questions. The first set of prompts aims to survey contacts that are of a certain regular frequency, in an emotional way important, or that provide help. The second question explicitly aims to survey contacts which ego is interacting with in his leisure time. The number of possible mentioned Alters is unlimited. The instrument surveys a high amount of non-kin contacts. That fact is correlated to the explicit request to list leisure contacts.

The general interest of transport planning in social contacts that have an influence on personal travel behaviour make it necessary to use a generator that is connected to the themes of transportation and travel. For that reason, another instrument was developed at the IVT. This instrument is similar to the Fischer approach. Here, nine stimuli are given, each of them in a separate question.<sup>1</sup> Two questions deal with in the past planned mutual visits to the living places of Ego or Alter, five questions include a stimulus for a scheduled leisure time meeting, one a scheduled work meeting. The last question asks for persons that are relevant for Ego but not mentioned in the other questions. Emphasising the planned or appointed character of a meeting is an attempt to distinguish the contacts of importance for each Egos' behaviour from random meetings with less important acquaintances. The number of possible mentioned Alters is unlimited. The instrument will be tested in a pre-test from June to September 2008 at IVT. The results can be summarised in the following way:

---

<sup>1</sup> Literally translated the questions are:

1. Which persons do you frequently invite to your home?
2. By whom are you frequently invited to their home?
3. With whom do you do errands (grocery shopping, other shopping, obtaining services)?
4. With whom do you participate in sports?
5. With whom do you engage in performing club obligations and assignments?
6. With whom do you attend cultural and leisure events (e.g. theater, concerts, etc.)?
7. With whom do you take vacations/holiday (e.g. short vacations, long vacation journeys, weekend visits, pilgrimages, etc.)?
8. With which individuals involved in your work do you take part in work activities (business trips, continuing education, training, meetings, conferences, errands for employer, etc.)?
9. Which persons you are talking about important things of privacy to or which you ask for advice in urgent affairs were not captured by the former questions?



Table 1 A comparism between the different name generators

Generator	Fischer (NCCS)	Burt (GSS)	Frei and Axhausen IVT 2006	IVT 2008
Amount of given stimuli	11	1	4	9
Limitations for the amount of possible given names	none	5	none	none
Survey instrument	questionnaire	questionnaire	questionnaire and supporting interview	questionnaire and supporting interview
Average size	18.5	3.0	12.4	11.3
Share relatives	0.4	0.6	0.3	tbd

Source: adapted from Frei and Axhausen (2007, 12) and own data

Every name generator technique surveys only a fragment of Ego's whole social network. While there is literature reporting contact numbers from 250 up to 5000 contacts for total networks, depending on the definition for social contact used (McCarthy, Bernard, Shelley and Johnson, 1997), the name generator provides a selection of the total number of contacts by stimulus. While the technique samples egos remembered and contacts mentioned and not the existing contacts of ego, there is danger for bias. In empirical work, it was recently shown that there is bias toward an observation of Alters with many contacts within Ego's network and towards Alters that are members of Ego's network but also provide contacts to other networks than Ego's (Marin, 2004). So the main handicap that goes hand in hand with the name generator technique is the selection bias arising from the fact that all gathered information are reported from Ego and so determined by his point of view. This can be a big problem, as no person knows every attribute of his friends, or even worse, how often these friends meet and interact without Ego. To handle this problem and verify Ego's information there is only one way. One has to ask Ego's contacts, too.

#### 4.1.2 The connection of ego-centered networks as an iterative survey

To connect ego networks in an iterative survey is a method that is called snowball sample. Such a snowball sample belongs to the chain methods and is a kind of an ascending survey sample. Here the questions of the survey are not only asked to ego but also to his alter, that themselves than become egos. A snowball sample is therefore characterised as „A technique for finding research subjects. One subject gives the researcher the name of another subject, who in turn provides the name of a third, and so on“ (Vogt, 2005, 300). An ascending survey method does not deal with a random sample taken at one point in time. The survey population is found by using hints from other persons (see van Meter, 1990). Snowball samples are mainly used in two situations. In most cases it is a method to provide access to hidden or hard to reach populations, e.g. drug users or persons with sexual diseases. Traditional methods of taking a sample do not work in such cases, particularly when the attribute of interest is very rare and implies the risk of social discrimination (Gabler, 1992). The second situation that makes the use of a snowball sample meaningful is to combine ego-centered networks to get

some kind of total sample of a given network. The information given by Ego is now combined with the information given by his Alters. So a detailed picture about Ego's network can emerge that is similar to the picture of the complete network of Ego's social contacts of interest.

#### **4.1.3 The Problem of Data generalizability**

In general, ascending survey methods imply a problem with regard to a possible generalizability of the data. The advantages and disadvantages of a snowball sample can be summarized in the following way: „However, the specific adaptation of ascending methodology is not obtained without the loss of easy generalizability. In exchange, descending methodology cannot reach hidden populations without specific adaptations“ (van Meter 1990, 32). Whilst drawing broad generalized conclusions from the data sample is only possible in the case of representative samples, the possibility of doing this is limited in the case of snowball samples. Because of the mentioned problem that Ego is giving the name of his contacts, and these Alters are therefore not selected by a random mechanism, only the initial sample, on iteration level zero, is a random sample. To continue with the sample after the initial level, also called level 1, on the other levels, that is iteration level 1 to iteration level n, means to include data that are arbitrarily or intentionally chosen by Ego and influenced by the network structure ego is embedded in. „Chains reflect aspects of network structure because such structures limit the possible paths for chains to take. Of all possible paths, the ones actually traced by a sample of respondents depend in part on their decisions about sending chains onward – for example, their decisions to pass rumors to certain acquaintances but not to others“ (Erickson, 1979, 277). This limitation regarding generalizability clearly emphasises the importance of the initial random sample. Comparing the results of this iteration-zero level with the results of the levels from iteration 1 to n can provide information about the magnitude of the survey bias.

Beside all these provisos, data resulting from snowball sampling strategies are not useless at all. The method is appropriate to survey aggregate data. Besides it allows verifying the information given by ego. It enables researchers to discover indirect relations between the actors of a network by combining multiple points of view. Drawing conclusions is possible for (Erickson, 1979):

- The individual actors and their attributes. This is important for computing of the degree of homophily between two actors within one network.
- The chain of contacts. Possible computations deal for example with their average length. Also the size of contact loops can be discovered.
- The chaining process. It may be interesting to compare the names given by Ego towards a specific question of the name generator und the names given by an Alter to the same question. By doing so it becomes observable if Ego and Alter acknowledge one another as important towards the same stimulus of the name generator.
- The whole network. By collecting direct and indirect relations from a multiple point of view a more precise network structure is sampled, compared to the non-iterative sampling of ego networks.

Bias can especially result from the following aspects of iterative ego-centred network samplings (see Atkinson and Flint, 2001; Berg, 2006; Erickson, 1979; Pool and Kochen 1978):

- Bias towards more cooperative persons. This form of bias is a problem for most survey methods.
- Bias towards persons that share many contacts. For these persons a higher likelihood is given to become part of the sample than for persons with few contacts or even isolated ones. Therefore persons that share many contacts are going to be over-sampled. This bias is also called selection bias.

- Information to provide contact can be denied by Ego or only be given in parts. If this happens in a systematic way, maybe because Ego does not want a specific Alter to participate in the study, it is called gatekeeper bias. It is of special interest because it may result in a missing chance to contact a specific Alter.
- In general mutual contacts are reported more often. Also those interactions that directed upwards in the social structure are mentioned more often than those that are directed downwards. Beside this it is of high importance that especially the questions of the name generator are formulated in a clear way. The more room for interpretation is given here, the more the structure of the network is going to have an effect on ego's answers.

Summarizing the problems of a snowball sample, Erickson says: „The pragmatic applicability of snowball sampling is limited by the need for fairly cooperative subjects, the advisability of unambiguous relational questions, and the possible confusion between chaining processes and structural effects. In addition, the method is hard to apply to relationships of which the respondent may have a great many, for example, weak ties. Finally, it is impractical to use more than a handful of waves lest nonresponse (cost aside) get severe, so that the overall structure of large networks is difficult to assess“ (Erickson, 1979, 288).

Approaches to reduce bias mainly deal with two topics. The first is to maximize the response rate. The resulting data basis is more substantive and contains on average less bias. For this reason it is advisable to start the chains from many different nodes. Also rewards, given to the respondents may help to increase response rates (see Dillman, 2000; Heckathorn, 1997). The other way of handling bias is to limit its effects by weighting the data. Colleagues of the group of Transport Systems Planning and Transport Telematics from the Institute for Sea- and Land-Transport of the Technical University of Berlin, which is also involved in this project, are working on that problem.

#### **4.2 The combination of sampling ego networks and activity related diaries**

The questions in the name generator and the following name interpreter, the part of the questionnaire that asks Ego to provide information about the socio demographic characteristics of his Alters, to specify the kind of his relation to these Alters and to report which Alters are, from his point of view, connected in a group or clique, aim to provide data about the network structure and characteristics of the participating persons. The interest of transport planning in the contacts of a person which influence travel behaviour consists of interest in the share of actions that are accomplished with others. Also it would be interesting to know which person in a given network plans such interactions. The traditional name generator technique is not appropriated to answer this question. „With the possible exception of personal network studies of interactions within bounded groups or organisations, researchers with a clear interest in „daily contacts“ and routine interactions may be better served by passing over name generators for diaries, or other procedures where respondents report on interactions shortly after they occur“ (Marin and Hampton, 2007, 167).

As the questioning technique is very common in the field of transport planning, it was not difficult to design an activity-related diary for the pre-test. The used instrument is a further development of a well tested and often used instrument. It is based on the Thurgau and MobiDrive diaries (see Axhausen, Löchl, Schlich, Buhl and Widmer, 2007). The original diaries were enlarged for the purpose of the planned survey. New are questions that ask which persons have shared the reported activity, if the activity was performed for the first time, if it was performed several times before or if its performance is routine and which person planned it.

It is planned that the respondents report their activities for eight days in a sequence. To survey one more day than a week's duration should provide information about the average number of routine activities returning in a weekly rhythm. The pre-test version is a printed questionnaire

once again. It is planned to use an online questionnaire and diary for the main survey. To use an online survey would provide the option to access the names reported in the name generator per drag and drop, so that respondents do not have to write down the same names again and again. Working with this option means therefore reducing respondent burden and improving data quality.

## **5. The survey process:**

Both the ego network questionnaire and the activity related diary are designed in a way that is strongly influenced by the suggestions of the Tailored Design Method of Don Dillman (Dillman, 2000). Also, the strategy of providing contact to the frame population follows that method. Dillman describes successful methods to reach a high respondent rate and to collect high quality data. His suggests:

- A strategy to open up and achieve the contact between the researcher and the respondents.
- A survey instrument build up in a logical way that contains minimal respond burden as possible.
- Accompanying measures to establish trust und increase the participation rate.

Summarised, Dillmans aim “(...) is the development of survey procedures that create respondent trust and perceptions of increased rewards and reduced costs for being a respondent, that take into account features of the survey situation, and that have as their goal the overall reduction of survey errors“ (Dillman, 2000, 4).

### **5.1 Population of inference, frame, and survey population**

The starting point of the planned iterative survey is the canton of Zurich. As mentioned before, the initial sample on iteration level zero is of special importance for the survey as it does not include bias that will hardly be avoidable on the following iteration levels. To be able to compute the dimension of this bias, the initial sample provides the basis for comparisons with subsequent iterations. Canton Zurich seems to be a representative area at least for Switzerland, as it contains nearly all forms of social life. Two of the biggest cities of Switzerland, Zurich and Winterthur, can be found here, as well as small towns and rural areas.<sup>2</sup>

Of interest for the survey are all adult persons living in this area. Because of the highly sensitive data resulting from the name generator and the name interpreter, children and teenagers under 18 years are excluded from the frame population. Of all persons of interest, only such can be surveyed which have a postal address and a phone number that can be found in the phonebook. The pre-test has taken place in Zurich City. For this purpose 150 random addresses and phone numbers were bought from an address dealer.

### **5.2 Opening up and making the contact**

The contact is initiated by a prenotice letter. It contains short information about the background of the study, the interest of transport planning in social network theory and what the questions are going to be about. An explicit advice about the iterative method of the survey is included and that respondents are asked for the full names of their social contacts and the postal addresses of this persons. Because of these sensitive items data protection and the respecting of privacy are assured by clarifying that this data are only needed to ask the persons for a participation in the study, too and making anonymous statements for example about the average distances between the living places of persons. The letter closes by asking the respondents for help to develop a new way of explaining traffic problems and stressing out how important each answered questionnaire is.

Three to four days after this letter, a phone contact is initiated to establish the person's willingness to participate in the study. If a person is willing to take part, the interviewer asks for an appointment to help the respondent with its' work in form of an interview. The questionnaire is sent to the respondent within the next days. It is accompanied by a coverletter that asks for help and assures data protection once again.

---

<sup>2</sup> To get an idea about the distribution of the socio demographics in the canton of Zurich see <http://www.statistik.zh.ch/raum/index.php?p=3>, August 2008.

Depending whether there is an appointment for an interview or not, the fourth contact is the interview or the questionnaire mailed back by the respondent. Particularly, with regard to the data quality and the importance of the initial level, iteration zero, it seems to be good to achieve an appointment for an interview for most respondents of this level. The interviewer can help in the case of problems and, especially for the questions of the name generator, can dig deeper. That is of importance as most respondents then mention further names (Wolf, 2004).

After the questionnaire has arrived at the IVT a thank you letter is send to the respondent. It accents the importance of the participation. It is also designed as a prenotice letter for the activity based diary. The importance to participate in both parts of the study is explicitly mentioned.

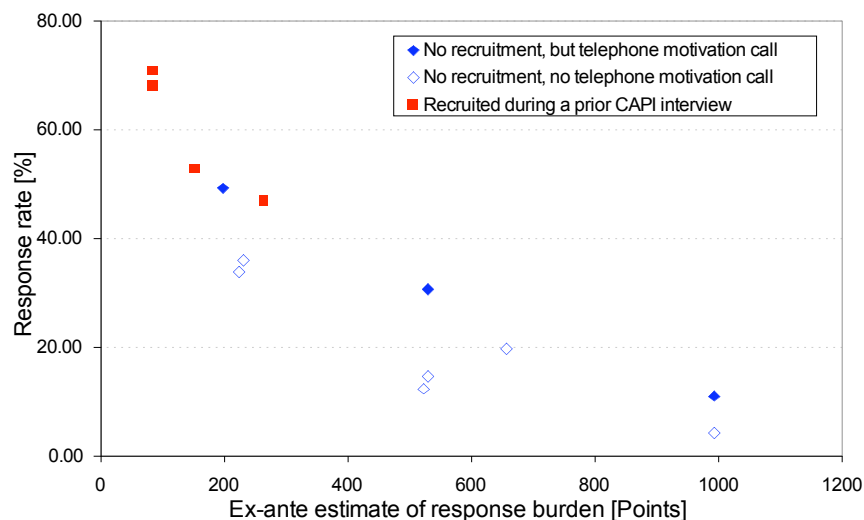
The sixth contact is made by phone once again. It aims to explore the willingness for participation in the diary. If a respondent affirms his willingness, the diary is mailed to him. The following steps are identical to the ones in the first part questionnaire. After the mailing back of the diary, a final thank you letter is sent.

To stay in contact in this frequent and regular way has the aim to reduce nonresponse. Giving the respondents a feeling of personal efforts for their participation has a positive effect and increases the respondent rate (Dillman, 2000).

### 5.3 The link between respondent rate and respondent burden

Every question is associated with an effort for the respondent to answer it. The sum of these efforts is the response burden. To make questionnaires with respect to this burden comparable, Axhausen 2007 describes a rating system used by a commercial survey research firm (Axhausen, 2007). Applying this method to former questionnaires used at the IVT, he discovers a nearly linear correlation between response burden and response rate. This can be used to get an idea about the response rates of future surveys.

Figure 2 Response burden and response rate



Source: Axhausen (2007, 3)

To compute the response burden of the first part of the survey, the ego network questionnaire is differentiated into four parts. Part one elicits the socio economic attributes of Ego. The

second part is the name generator. Here the burden for both instruments, the IVT generator from 2006 and that from 2008, is computed. The burden of the name interpreter is calculated for 12 alter, as that is the average number of names given to the IVT instrument used in 2006. The last component contains questions about the structure of the network. Ego is asked to arrange his contacts in groups or cliques. The burden for this component is calculated for 12 Alters, too. All in all the result for the questionnaire is around 730 points, corresponding to a response rate around 20%. Because of the higher number of questions the new generator implies a larger response burden. The burden implied in the diary is not calculated as the paper and pencil version will be modified strongly when implemented in the online survey. It can be assumed that the response rate will lie far under the rate of the part one questionnaire.

#### **5.4 Establishing trust to increase the response rate**

The use of a survey method that needs very sensitive information, names and postal addresses, to be supplied makes it necessary to intensify the efforts to establish trust beside the usual measures like establishing close contact or using the name of the institution to show the scientific background of the survey. Here three complementary measures are used.

The first one is an internet page to provide further information about the study background. Also, information about both parts of the survey and the purpose of the questioning is provided. Besides that, all persons of the research team are introduced and related former work of the IVT is linked.

To let the contacts of subsequent waves, iteration 1 to n, know that they are not the first respondents a postcard was designed. After asking Ego the questions about his network, he is given the opportunity, if he wants, to use the cards to prepare his contacts for the arriving prenotice letter from ETH. This measure will hopefully have a significant effect on the response rate. „Knowing that other people like themselves have completed a similar action can strongly influence people’s willingness to comply with a request“ (Dillman, 2000, 17). Furthermore respondents recognise that people they are trusting have taken part in the survey. The third method to increase the response rate is to provide a 20 Franks reward. This gratification is mailed together with the questionnaire. By sending it on this early point of the contact chain, it helps to establish trust. Respondents recognise that the research team is trusting them as they need not mail back the filled-out instrument, but could take the money and throw away the questionnaire. Giving the money to a person who has not done an activity yet provides some kind of social exchange situation. Respondents could easily adduce their part by responding the questions and mailing back the survey. „Much research has shown that „token“ incentives given with the request to complete a questionnaire, a form of social exchange, consistently improve rates. (...) However, a promise to pay people for completing a questionnaire by sending them a payment afterwards (economic exchange) does not“ (Dillman, 2000, 14f.). But to use such a method also implies the possibility of data bias. It must be proved if the 20 franks offered respondents are in a systematic way different from those respondents who did not know anything about a gratification. To achieve this comparison the initial level respondents will not know anything about the gratification strategy. They will therefore get it after they have sent back the filled out questionnaire. In this way an unbiased basis is sampled that can be compared to the later respondents.

Another way to increase the response rate follows a strategy of secondary gratification. That is to provide a payment for every recruited Alter that Ego has mentioned to the name generator and towards he has made efforts to get this person part of the sample (Heckathorn, 1997). This measure was not tested yet.

## 6. Results from the Pre-Test

Here first results from the pretest are presented. The results are only interim, as this phase is not completed yet. The current status of the dataset as of 11 August, 2008 is presented in the following tables:

Table 2 Response Rate

Phase	Pre-test			Of which Initial level		Of which Iteration 1 level (Alters from 3 out of 8 respondents on initial level)	
	overall	contingent (%)	realised phone contacts (%)	overall	realised phone contacts (%)	overall	realised phone contacts (%)
Sample	189	100.0		150		39	
Incomplete addresses	21	11.1		12		9	
Not reached by phone	86	45.5		80		6	
Reached by phone	82	43.4	100.0	58	100.0	24	100.0
Recruited	32	16.9	39.0	24	41.4	8	33.3
Participated	11	5.8	13.4	8	13.8	3	12.5
With interview	5	2.6	45.5	5	62.5	-	-
Without interview	6	3.2	54.5	3	37.5	3	100

Table 3 Name generators in the case of a snowball sample: A comparison

Generator	IVT 2006			IVT 2008		
	overall	initial level	iteration 1	overall	initial level	iteration 1
Returned questionnaires	8	6	2	3	2	1
Not yet returned questionnaires	8	6	2	13	10	3
Contacts named $\emptyset$	17.8	11.8	35.5	11.3	10.0	14.0
Minimum number	3	3	31	6	6	14
Maximum number	40	25	40	14	14	14
Range	37	22	9	8	8	-



## **7. Perspective**

Further analysis will be done when the pre-test is complete. Especially the effect of the rewards and the use of the greeting postcards will be evaluated with respect to the participation rate. Also results and the participation rate of the activity diary will be analysed. As the test includes two levels of the snowball, the initial and iteration 1 level, and all respondents will be asked to participate in the ego network questionnaire and the activity diary, the test can be expected to be complete in the mid of September.

Beside these statistics both parts of the survey project, the ego-centred network questionnaire and the activity diary will be turned into an electronic version to be able to transmit it per e-mail. It is expected that the response rate will drop when using the electronic instruments. For example it is planned that a given name of a social contact that is of importance for more than one stimuli of the name generator must be written down only one time and can be chosen from a list of all Alters names given by that Ego afterwards. Another advantage will be the use of google maps to geo-code living and work places. Using this feature respondents do not have to look for addresses but can mark the places of interest in a map. After constructing the electronic instrument there will be a further pre-test. This test is planned for the second half of October.

## 8. Literature

- Atkinson, R., and J. Flint (2001) Accessing Hidden and Hard-to-Reach Populations: Snowball Research Strategies, *Social Research Update*, **33**, University of Surrey.
- Axhausen, K.W. (2005) Social Networks and Travel: Some Hypotheses, in Donaghy, K.P., Poppelreuter, S. and Georg Rudinger (ed.) *Social Dimensions of Sustainable Transport. Transatlantic Perspectives*, 90-108, Ashgate, Aldershot.
- Axhausen, K.W. (2007) Predicting response rate: A natural experiment, *Arbeitsberichte Verkehrs- und Raumplanung*, **434**, IVT, ETH Zürich, Zürich.
- Axhausen, K.W., M. Löchl, R. Schlich, T. Buhl and P. Widmer (2007) Fatigue in long-duration travel diaries, *Transportation*, **34** (2) 143-160.
- Axhausen, K.W. (2008) Social networks, mobility biographies and travel: Survey challenges, *Arbeitsberichte Verkehrs- und Raumplanung*, **343**, IVT, ETH Zürich, Zürich.
- Berg, S. (2006) Snowball Sampling-I, *Encyclopedia of Statistical Science*, **12**, 2<sup>nd</sup> ed., 7817-7821, Wiley & Sons, Hoboken, New Jersey.
- Burt, R.S. (1984) Network items should be included in the General Social Survey, *Social Networks*, **6** (4) 293-339.
- Carrasco, J.A., B. Hogan, B. Wellman and E.J. Miller (2007) Collecting social network data to study social activity-travel behavior: a egocentric approach, *Environmet and Planning B: Planning and Design*, advance online publication, 1-21.
- Dillman, Don A. (2000) *Mail and Internet Surveys. The Tailored Design Method*, 2<sup>nd</sup> ed, Wiley & Sons, New York, Chichester, Weinheim.
- Erickson, B.H. (1979) Some problems of inference from chain data, *Sociological Methodology*, **11** 276-302.
- Fischer, C.S. (1982) What do we mean by "Friend"? An inductive study, *Social Networks*, **3** (4) 287-306.
- Frei, A. and K.W. Axhausen (2007) Size and structure of social network geographies, *Arbeitsberichte Verkehrs- und Raumplanung*, **439**, IVT, ETH Zürich, Zürich.
- Gabler, S. (1992) Schneeballverfahren und verwandte Stichprobendesigns, *ZUMA-Nachrichten*, **31**, Jg. 16, 47 - 69.
- Heckathorn, D. (1997) Respondent-driven sampling: A new approach to the study of hidden populations, *Social Problems*, **44** (2) 174-199.
- Jansen, D. (2003) *Einführung in die Netzwerkanalyse. Grundlagen, Methoden Forschungsbeispiele*, 2. erw. Aufl., Leske + Budrich, Opladen.
- Marin, Alexandra (2004) Are respondents more like to list alters with certain characteristics? Implications for name generator data, *Social Networks*, **26** (4) 289-307.
- Marin, A. and Keith N. Hampton (2007) Simplifying the personal network name generator: Alternatives to traditional multiple and single name generators, *Field Methods*, **19** (2) 163-193.

- McCarty, C., HR. Bernard, GA. Shelley and EC. Johnson (1997) Eliciting representative samples of personal networks, *Social Networks*, **19** (4) 303-323.
- Pool, I. and Manfred Kochen (1978) Contacts and Influence, *Social Networks*, **1** (1) 5-51.
- Wellman, B. (1996) Are personal communities local? A Dumptarian reconsideration, *Social Networks*, **18** (4) 347-354.
- Wolf, C. (2004) Egozentrierte Netzwerke. Erhebungsverfahren und Datenqualität, *Kölner Zeitschrift für Soziologie*, **Sonderheft 44**, 244-273.