Modelling social aspects of travel behaviour
A preliminary review

Nicole Ronald, Theo Arentze and Harry Timmermans
Urban Planning Group, Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven
The Netherlands
n.a.ronald@tue.nl

Key words: agent-based models, social networks, travel behaviour

Abstract:

Recent travel forecasting models have focussed strongly upon the fact that travel is derived from the activities in which people participate, such as work, school, shopping, sport, leisure, and social events. Participation in social activities is determined by one's friends and the groups that one is a member of, i.e., their household, their workplace/school, sporting groups, voluntary organisations and clubs. These acquaintances form part of an individual's social network: a representation of the people one interacts with. This paper presents a review of the intersection between social networks and travel behaviour, identifies the key concepts, and discusses how agent-based models could be used to explore the effects of social networks on travel behaviour.
1. INTRODUCTION

Recent travel forecasting models have focussed strongly upon the fact that travel is derived from the activities in which people participate, such as work, school, shopping, sport, leisure, and social events. Non-discretionary activities such as work and school can be partly explained by the traveller’s sociodemographic characteristics and generalised travel costs (Hackney and Marchal, 2007), as well as long-term decisions such as a decision to move to a particular town.

Participation in, and scheduling of, other activities is not as easily predicted. Participation in social activities is determined by one's friends and the groups that one is a member of, i.e., their household, their workplace/school, sporting groups, voluntary organisations and clubs. These groups form part of an individual's social network. This network is a representation of the people one interacts with. These networks often also contain an indication of how people are connected and how measure of the strength of that connection.

The main processes associated with social networks are selection and influence. Networks evolve over time, as people leave jobs, move, join new clubs, and meet new people. The strength of the connections between people also change over time. Interaction with others is required to maintain relationships. These activities may not necessarily involve travel and meeting with people in person, especially in the current environment where email, phone, instant messaging, and social network sites are widely accessible. However, we are mainly interested in day-to-day/short-term social activities that do have a travel component.

Individuals influence each other by sharing information, both intentionally and unintentionally. Relevant examples include telling a colleague about a new location, which may then become an option in the colleague's choice set for future activity locations with other friends. As such, these exchanges influence the dynamics of travel behaviour.

Understanding the social network that lies on top of the spatial network could lead to better prediction of social activity schedules and therefore better forecasts of travel patterns, in particular for social and leisure activities.

Our aim therefore is to develop an agent-based simulation model to investigate the effect of social processes on activity and travel behaviour. The model will be an experimental tool. This paper presents a review of relevant theory and discuss several issues with modelling the effects of social networks on travel behaviour within a preliminary framework. Background material is presented on the evolution of transport models and social networks, followed by a more detailed discussion of the key concepts.
Several existing models are described, followed by some notes on implementation and future work.

2. BACKGROUND

2.1 Transport modelling

Computational models of transport behaviour have evolved since their first use in the 1950s. This can be partially attributed to the increasing availability of high-performance, low-cost, computer hardware, as well as the need for more accurate forecasts and understanding.

The various types of models developed up to the present time can be categorised in three generations:
1. trip-based models (4SM);
2. disaggregate/discrete choice models (both analyse trips independently of other trips made by the same person);
3. activity-based models.

The most well-known approach for modelling transport is the “four step” model. This model was developed in the 1950s for looking at travel behaviour at an aggregate level. The network usually consists of zones, which can be based on existing well-defined zones (e.g., postcodes) or zones created by the modeller.

A strength of the four step model is its simplicity. Despite this, many aspects of the model have been criticised. According to Banister (2002), the process is too rigid and the model is more concerned with reducing individual travel time instead of other travel measurements. It was created to be simple, however this fails to recognise that transport is complex (Boyce, 2002). The model is based on trip-based methods, rather than activity-based methods (McNally, 2002), which contradicts what is commonly understood about travel behaviour in that is derived from activity. It also does not take into account individual choice (Golledge and Garling, 2004) and is not non-behavioural, in that it does not represent “the decisions that consumers make when confronted with alternative choices” (Domencich and Fadden, 1975).

In the late 1960s, there was a realisation that aggregate models did not meet the needs of planners (Ben-Akiva et. al., 1996). Researchers started to develop disaggregate models of travel behaviour and by the end of the 1970s some were operational. Travel demand is the result of decisions made by an individual (Bierlaire, 1998), therefore choice models appeared to be a suitable approach. These decisions could be about activity, destination,
departure time, mode and itinerary. The “unit” is still single trips, like the trip-based models. However, later models began to incorporate the notion of interrelated activities. The elements of choice models are the decision makers, a set of alternatives to choose from, the attributes of each alternative, and decision rules. Each decision maker may have a different set of alternatives.

The main drawback to the choice model approach to transport models is that the processes that result from policy or societal changes cannot be fully represented in a choice model. The relationship between activities across a day is also not an important consideration (Doherty and Ettema, 2006). Marchal and Nagel (2005) also discuss technical drawbacks, claiming that the use of utility models for large transport models is infeasible, as the number of choices is too great, and that the actual decision process is not modelled, just the outcome.

Activity-based models were proposed in response to the shortcomings of the trip-based model. These approaches “aim at predicting which activities are conducted where, when, for how long, with whom, the transport modes involved and ideally also the implied route decisions” (Arentze and Timmermans, 2000). Participation and allocation of activities are generally based around households.

Several sources (Axhausen, 2000; Kitamura, 1988; McNally, 2000; Wang, 1998) list elements of the activity-based approach. Some lists overlap with others. The main elements are:

- The key concept is that travel is derived from activity participation.
- Activities are influenced, planned, and executed in coordination with one's household and social networks.
- The focus is on sequences of activities rather than single trips or single activities.
- There are various constraints on activities.

Axhausen (2000) defines an activity as “a continuous interaction with the physical environment, a service or person, within the same socio-spatial environment, which is of importance to the person.” This definition includes the possibility of interacting with others, however limited to one other person.

The next shift in travel forecasting models is to then determine from where activity participation is derived. For fixed activities such as work and school, the frequency, location, and duration are clear. However for social activities, there is a lot more flexibility in frequency, duration, and location and are more difficult to predict (Doherty and Ettema, 2006).
The consensus so far is that social activities are driven by the members of one's social network, in particular “determining trip destination, frequency, mode and scheduling” of leisure and “personal” activities (Hackney and Axhausen, 2006). Buliung and Kanaroglou (2007) state that some researchers are already looking beyond households to the influence of social networks.

2.2 Social networks

Scott (1991) describes social networks as consisting of relations and attributes. Attributes are the attitudes, opinions and behaviours as properties of actors and groups. Relational data are the connections, ties, and contacts between the actors. He also describes many measurements for social network analysis such as cores, cliques, clans, and plexes (Scott, 1991).

There are two ways of looking at a network (Carrington et al., 1995). We can look at the global properties of a network, where the network is usually complete and closed (i.e., all connections between a pre-defined set of people are known). We can also analyse the network from the perspective of individual actors in the network, also known as an egocentric network. Focal actors (or egos) are sampled from a larger population and are asked about the actors in their networks (or alters). In some cases, only basic details are collected about alters, so that they remain anonymous. However, a snowball survey approach may be used, in which the alters are then contacted and asked to provide details of their alters. The egocentric viewpoint is more useful when the actor set is not completely defined. Egocentric networks are the most useful for our intended model, as the network will not be closed and therefore constructing a complete network is not possible. This echoes the discrete choice approach described earlier, which is also egocentric.

Much research has been done on the structure of whole networks and defining different structures and network analyses and measurements. Structures such as the small-world model (Watts and Strogatz, 1998) and the scale-free network (Barabasi and Albert, 1999) have received much attention from researchers. Unfortunately these structural analyses are more appropriate for whole networks.

The structure of the network is important, but how the structure changes over time is also important and this is not as well studied. Newman (2003) states that we need to understand the structure first before we can begin explaining the workings of these systems. The social networks we are interested in will change over time: the attributes of people may change, which may then lead to changes in the connections between people, which then may lead to more change in personal attributes, and so on.
Network processes are the “series of events that create, sustain, and dissolve network structures” (Doreian and Stockmen, 1997). Two types of network processes are defined by Robins et. al. (2001). The social selection process occurs when “individual attributes contribute to the formation or change of network ties”, whereas the social influence process occurs when actors are influenced by their alters and their attributes change. It has been postulated that the effects of these processes are difficult to distinguish. Leenders (in Doreian and Stokman, 1997) states that “actors will shape their networks and simultaneously, are influenced by the structure of the network”. The definitions of selection and influence are slightly different. Selection occurs when “actors consciously or subconsciously structure their network”. Influence is also referred to as contagion and occurs when “a social actor (partly) adapts his behavior, attitude, or beliefs to the behavior, attitudes, or beliefs of other actors in the social system”.

Newman also states that “typical social network studies address issues of centrality (which individuals are best connected to others or have most influence) and connectivity (whether and how individuals are connected to one another through the network)” (Newman, 2003). These two principles can also be described as influence and selection.

3. CONCEPTUAL FRAMEWORK

From the literature, it is clear that networks and network analysis is useful in a range of areas. However, there are still areas that are not well understood. Most work has been done for looking at structures in the natural and physical worlds, such as food webs/chains, power grids, and the internet amongst others (Strogatz, 2001), and defining different types of networks and measurements of properties for these networks. However, Boguna et. al. (2004) identify three network elements where these networks differ from social networks, therefore the theories and networks developed cannot be directly applied to social networks without some modification.

The system we want to model contains people interacting with each other by meeting and sharing information. The properties of individuals can change, such as car ownership status or work/home location, leading to a different set of activity and travel choices. This may then influence their connections: by moving to a different city, they may lose contact with certain people and they may meet new people and undertake different activities with them. This points to dynamic networks and therefore the static theories of structure are not useful.

Arentze and Timmermans (2006) describe a framework and a small model where social activities are incorporated into activity-travel models.
Their aim is to describe “a modelling framework about the relationship between social networks and interactions that is particularly relevant for modelling spatial choice behaviour in the context of microsimulation systems”.

Two aspects to creation, deletion, and maintenance of ties are recognised: the need for the company of others (social needs) and the acquisition of knowledge (information needs).

The prototype model created looked at social needs only. Social connections between people have a “strength” value and are based on similarity on personal attributes (age, education etc.) and group membership.

The simulations described started with four groups of five people and all groups fully connected. Each day agents consider sending out invitations for interaction and may also receive invites. Interactions were found to be irregular but stable over time.

An assumption of the model is that similarity in attributes, preferences and action spaces increases the probability of a tie being created. After the creation of a tie, influence then leads to knowledge acquisition and attribute adjustment. They conclude that social networks and activity-travel patterns tend to co-evolve.

From this experimental model, we have identified three issues that will be important in developing a larger model:

- the selection dynamics and formulation of the social network;
- the influence processes in the networks: information exchange, social learning and social adaptation; and
- the activity/travel generation that emerges from the network: focussing on predicting the participants in, frequencies of, and locations of social activities.

### 3.1 Selection

The selection processes in our system are based on the formation and deletion of connections. When people travel somewhere, they may meet someone new at that location and make a new connection. In the process of life events such as moving, starting a new job, or joining a club, connections are created. If people do not communicate with an acquaintance often enough, the connection may eventually fade and disappear. We are interested in the reasons for the formation and deletion of connections and how this can be modelled.

Hackney and Axhausen (2006) claim that social networks can be generated using behavioural tendencies from sociology, including homophily, bridging social capital (where people are similar in one way but
different in another), and putting limits on the number of relationships. The latter property is one of the principles used by Jin et. al. (2001) in their investigation into growing networks. Other principles included increased chances of meeting another person if you and them have a mutual friend, and decaying tie strengths decay over time. Liben-Nowell and Kleinberg (2007) also examine the prediction of new ties and review several scoring measures, including neighbourhood measures (e.g., common neighbours and weighted variations) and path-based measures.

In their review of homophily, McPherson et. al. (2001) define homophily as the “principle that contact between similar people occurs at a higher rate than among dissimilar people”. The concept of homophily is credited to Lazarsfeld and Merton's article on friendship processes (McPherson et. al., 2001). Some of the attributes that are used as similarity measures include gender, age, education, occupation, social class, behaviour, attitudes, abilities, beliefs, and aspirations.

“In diverse societies, race, and race-like ethnicity create the most stark divides. Sex, age, religion and education also strongly structure our relationship with others. Occupation, network position, behaviors, and interpersonal values also show considerable homophily, but they seem to be more specific to certain types of networks and/or derived from the basic facts of sociodemographic homophily.” (McPherson et. al., 2001)

Distance is also a key factor in the maintenance of relationships. McPherson et. al. (2001) claim that the most basic source of homophily is space as we “are more likely to have contact with those who are closer to us in geographic distance than those who are distant.” Wong et. al. (2006) concur, although they specifically state it is a baseline homophily. An hypothesis put forward by Silvis et. al. (2006) is that people will make higher cost (time) trips for greater social payoff, i.e., to see more people at once, to see people that one has known for longer, and to see family. This was not completely verified by their investigation.

Van de Bunt et. al. (1999) describe a friendship “lifecycle” as moving from unknown, to neutral, then to either friendly or troubled. From friendly, the relationship can also become friend, and, possibly, best friend. Some reasons for dissolving ties are that interests, values and opinions are not shared and that maintenance costs may be too high (e.g., distance).

Zeggelink (1997) investigated several different actor-oriented models of network evolution. The models contained rational actors with characteristics and behaviours who were trying to optimise a goal function (maximising utility and minimising tension) by establishing, maintaining and dissolving relationships. In the friendship emergence model, the agents had three goals:
to have a certain number of friends, to have similar friends, and to belong to subgroups. Tension calculations are based on minimising the difference between what one has and what one wants with respect to these goals.

### 3.2 Influence

In our system, people influence each other by providing information or observing behaviour. For example, one of your acquaintances may inform you about a new location they have recently discovered and you may add it to your choice set for future use. Other indirect influences on travel behaviour include moving to be closer to work or family and choice of vehicle.

The majority of the literature on influence relates to investigating the spread of epidemics or rumours, the diffusion of innovations, and determining which nodes or individual in the network are “powerful”. These can be divided into two different objectives: where the property being spread is undesirable and needs to be controlled or eradicated (such as epidemics), and where the property being spread is desirable and should be spread quickly (such as rumours, new innovations) (Boccaletti et. al., 2006).

Valente (1995) covers a range of different aspects of diffusion models for innovations. He reviews both positional (influenced by neighbours only) and structural (influenced by the wider network) models. Two generic models of diffusion are described by Kempe et. al. (2003). The Linear Threshold Model starts with a random number of “activated” actors (i.e., those who have adopted the innovation) and each actor has a threshold. Once the sum of their neighbours' influences is above the threshold, the actor is then activated. The Independent Cascade Model also starts with a random number of activated actors, however each actor gets a once-only chance to activate its neighbours with probability $p$. The focus is on which actors to target for maximum influence.

In an interesting combination of epidemics and transport, Eubank et. al. (2004) presents a method for generating the network for determining the spread of disease using travel data. They use TRIPS to generate people's travel activity for a day. From that data, a bipartite graph linking people and locations is created and an indication of colocations and possibilities for spread is found. The use of bipartite graphs is commonly used for creating social graphs where there is some form of collaboration involved, such as on academic papers. Lauw et. al. (2005) also use bipartite graphs in their model of creating social networks based on online “colocations”, i.e., when two people visit the same website within a certain time window.
3.3 Activity/travel generation

Activities are generated due to “physiological, psychological and economical needs” (Wen and Koppelman, 2000). The different activities can be categorised as subsistence (work-related), maintenance (keeping the household running), and leisure.

There exists two schools of thought on activity generation for travel behaviour purposes. The first is that activities stem from desires. Chapin (1974) considered the propensity to participate in an activity along with the available opportunities. The other approach is that the activity set is constrained in various ways. Hagerstrand's (1970) constraints included capability (need to sleep and eat), coupling (need to coordinate with other people or comply with opening hours of institutions), and authority (need to conform to societal laws and customs) constraints. The former, more needs-based school is relevant for our work. More specific theories of determining duration and frequency are reviewed by Arentze and Timmermans (2000).

As Newman recognised, research has been slow in understanding the actual workings of networked systems (Newman, 2003) and the focus has been on structural form and analysis.

We are interested in how the activities are generated if a social aspect is included. In order to predict the travel undertaken as part of an activity, we require the participants in, frequencies of, and locations of social activities.

4. RELATED WORK

Some work has been done with trying to integrate the effects of social networks and interactions with activity patterns. This is very recent, as Axhausen (2006) notes that “transport research, but also sociological research has in the past not looked at the link between social networks, locational choices and travel”. Currently, no model thoroughly incorporates all three concepts we have identified. Most models focus on one or two concepts and in most cases at a very basic level. For example, initial social networks usually follow a generic random graph model, which is not realistic. However, most of these models are described as preliminary or proof-of-concept.

Sunitiyoso et. al. (2006) have investigated mostly influence, by exploring the spread of soft (or psychological) policy measures, such as environmental awareness and encouraging car-sharing, and in particular the influence of a minority influence group. The model also included meeting and communicating with other agents. Their experiments showed that diffusion did occur and that club membership was more effective at spreading
Modelling social aspects of travel behaviour

11

information than neighbours. Sunitiyoso and Matsumoto (2007) also investigated the spreading of mode choice behaviour. The model has two layers: a traveller model that models the decision making process, and a transport model that models the transport system and provides generalised costs as feedback. Agents can change their mode using two rules: those using the payoff rule find the neighbour with the best payoff from the last round and changes to their behaviour, and the conformists change their behaviour to match the majority of their neighbours.

In terms of activity generation, Carrasco and Miller (2007) developed a “proof-of-principle” model to look at including the social dimension in forecasting models. The focus was on how social activities can be generated from a social network. A key idea from this work is that the “potential activity and travel” between two connected people is more relevant to predicting travel behaviour than the tie strength alone.

Ettema et al. (2007) described some initial work on developing social network-explicit land use and transportation models. As a starting point for investigating how social interactions influence short- and long-term decisions, several forms of interaction are identified: task allocation, joint activity participation, help or support generation, and social externalities. The proposed framework is utility-based and consists of the sum of several utilities representing the effects of interactions from the points of view of who conducts the activity and who is affected or targeted.

Hackney and Axhausen (2005; 2006) also developed a simple model looking at how activities are generated. The agents are located on a grid and are provided with an amount of travel budget. At each time step each agent decides which of its neighbours (those that can be reached within budget) it would like to visit in order to increase its social status. All trips are home-based, so the agent needs to have enough budget for a return trip. The social status indicator is (normalised) betweenness centrality: the number of shortest paths that pass through the agent. As expected, those with better access to others had the higher social statuses at the end of the simulation.

Focussing on both influence and selection, the model created by Marchal and Nagel investigates where “individuals perform activities such as shopping and leisure” (Marchal and Nagel, 2005). Agents in the model have limited information about the environment and are connected to several other agents through a social network. Agents are provided with plans and travel around a network carrying out activities individually. Links with other agents are created or reinforced when agents travel to the same location and decay over time if a meeting has not occurred for a while. Each agent has some knowledge of the area around their homes and workplaces, as well as two buffers of “useful” locations and not so useful locations. Every timestep each agent randomly selects a cell that they are aware of and informs their
friends of this cell. The friends evaluate the cell and if the inclusion of that
cell into their travel plans leads to an improvement, then the cell is added
into their useful buffer of cells. If not, then it randomly replaces a useless
cell in their memory. This leads to some cooperative behaviour as the agents
are storing and possibly sharing information that is not useful for
themselves.

The essence of this model has been used as a basis for a recent extension
for MATSim (matsim.org, 2007) that allows for the inclusion of social
network data into a large-scale model (Hackney and Marchal, 2007). In this
model, individuals make visits to locations, alter the strength and existence
of ties, exchange information with each other, and modify their plans by
updating the location of secondary activities during each run. At the
moment, the selection and influence strategies are simple, but more realistic
behaviours could be incorporated.

With respect to simulation properties, both Marchal and Nagel (2005)
and Hackney and Marchal (2007) report on the computational aspects of
their models and make estimates of the computational complexity.
Comments are also made on the usefulness and practicality of exploring
social aspects and their effects (Ettema et. al., 2007; Hackney and Axhausen,
2006). Collection of more data will be required and it may be that
simulations with more data input and more detail do not provide an
improvement in forecasting. As a large amount of travel is now for
social/leisure purposes, it seems reasonable that we attempt to understand the
reasons behind it. Hackney and Axhausen (2006) are concerned that the
emergence of the social network may not be realistic.

5. IMPLEMENTATION IDEAS

Agent-based modelling is frequently used for applications where the
behaviour and intentions of heterogeneous individuals and interaction
between individuals is required. As presented in section 2, transport models
have evolved from aggregate models to individual-based models focussing
on behaviour, which appears to fit with the agent paradigm.

Both Bonabeau (2002) and Macal and North (2006) present lists of
system attributes that are ideal for selecting agent-based modelling for that
system, including amongst others:

- agents have dynamic relationships with other agents;
- relationships form and dissolve;
- agents have a spatial component to their behaviours and interactions;
- the topology of the interactions is heterogeneous and complex.
Our system consists of different people, their relationships and interactions with each other, and their movement around the transport system. The topology is not homogeneous and clusters may form.

Many packages for creating agent-based models currently exist. Some are based on particular agent types and some are more flexible, in that the modeller is allowed more freedom to create their own behaviours whilst taking advantage of a general framework handling time management and other aspects of the simulation. Many models are implemented without the assistance of a modelling package (i.e., implemented directly in Java or C#) as “many ABS tools and platforms make limiting assumptions regarding the way that entities are modelled” (Davidsson et. al., 2007).

We intend to use Repast Simphony (Repast, 2008) for implementing models. Repast is a reasonably mature agent-based modelling package which contains support for network relationships between agents, as well as providing flexibility to create our own behaviours in either Java or Groovy, a dynamic, script-like language that compiles to Java bytecode. Visualisation tools are also included.

6. CONCLUSION

This paper has presented motivations and issues involved in modelling the effects of social networks on travel behaviour. It is clear that there exists a relationship between one's social networks, activity choices and resulting travel and that this is a recent and interesting development in transport modelling.

The theory on social networks is vast and well-developed in some areas but not in others. The requirements of the system we are studying is that the network is dynamic and not closed (that is, egocentric), however the theory currently focuses on static and closed networks. Several existing transport models that incorporate social networks were described and related to our framework, however all are preliminary and concentrate on only one or two on the key concepts we have identified. The use of agent-based modelling appears to be a sensible approach for this model.

Our next step is to refine the framework using the various theories on selection and influence in the literature, followed by implementation in Repast and experimentation.

Modelling social aspects of travel behaviour


Kempe, D., J. Kleinberg, and E. Tardos, 2003, “Maximizing the spread of influence through a social network”, in: Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 137 – 146, ACM.


