Coevolving social and transportation networks

JK Hackney
Abstract

A component of work and non-work travel may be plausibly explained by the need for individuals to be co-present in order socially interact. Successful socializing requires maintaining and exploiting relationships, fulfilling obligations, and perhaps optimization of contacts. A social network is a way to map relationships and to quantify conduits of information, services, opportunity, and power. People must be able to navigate both their relationships and physical space in order to find other people and facilities which can provide the services they need. The result is that travel behavior and social structures affect each other. This research examines hypotheses of the interactions between social networks and trip generation. Models of social networks should be used by planners if their effects are significant and distinguishable from factors currently incorporated into transportation models.

A multi-agent-based network evolution model is presented in which social visits are generated by a set of attraction parameters which include agent attributes, travel impedance, and descriptors of the actors’ position in the social network topology. The trips are constrained by the agent’s travel budget, and dynamic equilibrium is maintained by removing or weakening old links. The destination choice is thus a tradeoff between cost and the need for co-presence, marked by correlations between actors in the changing social network. The interactions are monitored using a statistical toolbox from social network analysis and placed in the context of findings from network science literature on network evolution, as well as from empirical work in sociology on real social networks. The behavior model used for social networking is described, and results from an ensemble run are presented. A sample case of a transportation “improvement” is also presented, to illustrate the way in which the social network adapts to changes in transportation costs that ease spatial searches.

Keywords

social networks, social simulation, dynamic networks, multi-agent based modelling, spatial networks, activity space, social geography, network evolution

Preferred citation style

1. Summary and Research Questions

This paper presents an agent-based model that links a non-spatial social network with a transportation network. It is primarily an exploratory effort to characterize the dynamic dependence of social connections on travel impedance (cost, time, etc.) within a constrained budget allowance (time allocation) for travelling. The output is analyzed statistically.

Broad research questions include:

- What kind of social graphs emerge when agents’ travel is constrained?
- How do the topology of the transportation network and the spatial layout of the agents affect the social network?
- To what degree is non-heterogeneity observed in the social network, despite homogeneous agents, behavior rules, and transportation network?

The broad questions are formulated below in more precise hypotheses related to long-term behavior, land use and transportation.

Even though the model has been kept very simple, it has not been fully explored. The mixture of spatial and social networks introduces a very high dimensionality to the parameter space, and quantitative comparisons with other network growth models have not yet been completed. Qualitatively, a positive-skewed degree distribution is observed when travel cost is made non-homogeneous by reducing travel cost locally. This indicates social clustering typical of empirical findings, and model behavior consistent with recent network growth models based on other assumptions.

The behavior model used for social networking is described, and results from an ensemble run are presented. A sample case of a transportation “improvement” is also presented, to illustrate the way in which the social network adapts to changes in transportation costs, with a discussion in the context of hypotheses about spatial social networks from Axhausen (2005).

Results from verification are not shown here in detail, but the model does perform as expected in the region of the parameter space that was tested.
2. Framework

People travel for personal and professional reasons. It is important to understand the spatial and temporal patterns of travel in order to forecast demand for infrastructure and public transportation services, and to offer alternatives to mobility, such as land use planning, zoning, etc. Travel for work purposes or on major holidays is easier to understand and to model than travel for personal or recreational purposes, which is more irregular.

Transportation modelling limits its explanatory approaches at the moment to four levels: the generalized cost of the alternatives, the socio-demography of the travellers, their values, and their lifestyles. Random utility models and discrete choice analysis are usually applied to trip-based datasets to estimate mode, destination, or route choice (Ben Akiva and Lerman 1985, Train 2002). Behavior analysis and models use variables indicating family and household structure, mobility tools, place of residence, and the work or school location (Axhausen 2005). These more or less define the everyday travel. Such models can only capture a portion of the variability of travel behaviour (Stopher and Jones 2003, Lee and McNally 2003, Kurani and Kitamura 1996).

A new challenge lies in deepening the understanding and improving models of the growing private travel of individuals, especially as it occurs in the context of groups. The timing and spatial patterns of such travel is irregular and largely endogenous, as is the size of the travel group. Important contributors to traffic loads include the growing social practice of parents accompanying children to school, often with a car; retirees who are much healthier and more active than in previous generations, and who are more likely to still drive cars; the lunchtime travel of employees who have a vehicle available to them during the day; increasing flexibility in working hours enabling shopping during the day; later shopping hours in certain zones; and decreasing travel and communications costs which may be enabling people to maintain more far-flung social networks and encouraging them to travel more frequently and farther in order to socialize.

Two new related approaches seek improvements using agent simulation. Activity-based modelling derives an individual’s travel decisions from a model of a schedule of activities and attempts to capture trip chains (see overview in Mohammadian and Doherty 2004, Bradley and Vovsha 2004). This framework has permitted, among other things, the study of the influence of intra-household relationships on travel (Meister, Frick, Axhausen 2004, Srinivasan et al. 2005, Han, et al. 2005, Pribyl and Goulias).

While efforts at intra-household behavior models have been emphasized, models of the influence of inter-household social connections are in the conceptual stage (Joh et al 2004). The importance of the social network in social and large-space mobility has long been shown
Grieco 1988, 1996). Its role in everyday mobility, including the decisions regarding short vacations and excursions, remains undocumented, except for first estimates (Blinde and Schlich 2000, Roorda 1998). Dugundji and Gulyas (2003), Dugundji and Walker (2005) and Paez and Scott (2005) show that mean-field effects in a generated random social network have significant explanatory power for binary mode choice. Dugundji and Walker (2005) were able to improve mode choice models of work travel by replacing the traditional panel variables with the mean field influence.

A well-specified social network component in a travel model could improve the accuracy of trip destination and timing in the trip generation phase, and it could reduce computational time in the assignment (route choice) phase by narrowing the travellers’ choice sets according to what they would realistically know through their social contacts. A social network could also serve to model the spread of ideas about transportation alternatives or about new travel locations that would act as exogenous influences on travel behavior, specific to the social network. (Dugundji calls this difference the “identified” ego network, versus the “unidentified” alters, as in the mean field model for example).

A realistic first assumption in linking social and transportation networks is that maintaining relationships depends on regular visitation. It would therefore be reasonable to represent social networks as an emergent phenomenon dependent on travel opportunities, i.e. agents meeting with each other by travelling on the transportation network. The emergent social structures can then be examined with the knowledge that they involved very basic tradeoffs between the social desire (or need) to interact and the inconvenience of travel. This is the approach presented here.
3. Disadvantages of Social Network Analysis Approach

It has been argued (e.g. Professor M. Wegener, University of Dortmund, discussion during Frontiers of Transportation Conference 2005, Amsterdam) that one cannot know the “true” reasons for people choosing to travel at a certain time on a certain route, without asking them, and they will themselves seldom know the answer. The issue is important at the meta-model level when one is deciding what kind of model to use for travel demand. Without knowing the true reason that a decision is made, one cannot improve upon a model of the decision which is based on the observation of the circumstance surrounding the decision. In other words, building social networks into a transportation model may not help specify the model any more than existing proxies of behavior that we know already, like age, gender, income, job, vehicle ownership, time and type of day, and residential location. This argument has special relevance when one considers the significance in social network analysis of homophily (McPherson et al 2001): the notion that people associate with people who are like themselves. Since transportation models already take into account the observable characteristics of individuals, they may already implicitly model behavior that is based on social networks, without the additional computer or survey costs of explicitly including the social networks. Conventional engineering wisdom says if an improvement to the model doesn’t help arrive at a better result for cheaper, then it should not be included.
4. Agent Based Approach

The main argument against this point of view is that building social networks into travel models hasn’t been extensively tested, yet (for example Marchal and Nagel 2005, Dugundji and Walker 2005, Dugundji and Walker 2003, Paez and Scott 2005), and the costs and benefits of doing so have not been examined closely. An additional issue is also the definition of “the model result”. A current transportation model gives accurate results of traffic flows for certain circumstances. But it does not explain the travel behavior, that is, the reasons for initiating and sustaining travel and for doing it in the manner observed (Stopher and Jones 2003, Lee and McNally 2003, Kurani and Kitamura 1996). Reasons for behavior is not an output of current transportation models. Travel behavior itself is more an area of study than an application, but the arguments above about social networks imply that much in the way of more accurate traffic flows and destination choices may be gained by incorporating behavior into decision models.

It is true that little is known about how people make travel decisions because the observations are expensive and because so much of the behavior is habitual that people are not able themselves to say why they do what they do. An agent-based model allows the researcher to represent travellers as agents with behavioral rules and to see if realistic system-wide transportation behavior emerges, or grows, out of these rules. The result of the model in the sense of traffic flows may not be any different from the statistical model. And, the model might run slower than existing transportation models. But the agent approach allows a richer exploration of why the outcome might be occurring, and offers more possibility for experimenting with dynamic relationships between individuals, groups, and the transportation system. The particular agent behaviors might not be the actual behavior of people, but they permit hypothesis testing and they provide a basis for discussion of likely behaviors and for focusing further experimentation and surveying more efficiently. It is particularly important to use social networks for planning if nonlinear or complex effects emerge which would not conform to expectations of system-level models. In that case, a program that produces emergent travel behavior from agents would be a more appropriate tool for policy makers who may want to understand system-wide effects resulting from individual responses to policies.
5. Hypotheses

The topology (network shape) and spatial extent of social networks should correspond in some way to travel costs, residential location, and land use (location of activities). More complex social network characteristics should also be testable in a complete dynamic model, for instance:

1. The size of a person's "contact space" is inversely proportional to the generalized cost of travel.

2. With lower generalized transportation costs, people maintain contacts to more non- or partially overlapping networks (clusters).

3. Contacts are more selective with lower generalized transportation cost; social contacts are more valuable.

4. One is no longer constrained to be content with knowing only one's neighbors.

5. Networks are more valuable because less valuable relationships are dropped as innovation allows cheaper network expansion.

6. There is a relatively higher number of low-intensity contacts with lower transportation and communication costs (more weak ties).

7. Selectivity of contacts; one is able to pick more specialized relationship; more efficient consultations means fewer consultations.

8. The number of social activities an actor is involved in increases with his social capital.

One would hope to address such hypotheses eventually, though even a simple model might yield insights. For the time being, basic questions to pose include:

- How does the social network relate to accessibility?
- How do trip distance and frequency relate to the social status and social network?
- What distributions of social network statistics are observed?
- Are they related to the spatial distribution of the agents?

These questions will be discussed along with the hypotheses in the results section.
6. Model Description

The program is written in Java and uses the RePast (RePast 2003) and JUNG (JUNG 2003) libraries for agent, graphics, and network analysis and manipulation tools.

The program reads in the transportation network and the social network. The former is a non-directed, weighted topology, initialized to an orthogonal grid with identical link weights (costs). The latter is a directed graph with no weights (weight = 1 or 0), initialized to just a list of named agents with X,Y locations in space, and no social links at all. The program assigns the agents to nodes on the transportation network by matching the X,Y locations of the agents to the transportation nodes. There is no limit to the number of agents which can occupy a node, but agents should not be co-located until one has a special model for relationships when there is no distance between agents (see utility function). The transportation network and the social network are constructed by hand in Excel and saved in GraphML format (Ulrik Brandes 2005).

The agents attempt to maximize their social status, subject to a travel budget, by maintaining associations with other agents in a social network. Complex social behaviors like exploitation of social position or reciprocal obligations are not represented yet. Maintaining associations simply requires regular visitation (travel) to the location of the other agents. Associations which are not maintained can disappear randomly with increasing probability according to a logistic function of how long it has been since the last visit. The probability of removing a link is 50% when the link is between 4 and 5 time steps old (Table 1).
The agents are located on spatially distributed nodes between which movement has a cost. Each agent has a travel budget which it can spend in order to travel to visit other agents. Agents attempt to maximize their social status (power) within their reachable neighborhood, weighed against the inconvenience of travel, as described below. Neighborhoods overlap, so agents far from each other can affect each other over intermediate agents despite the travel budget constraint.

The travel budget is constant and equal for all agents and is renewed each time step. A list of neighbors is established in the setup phase for each agent using a Dijkstra shortest path search on the weighted transportation network. A neighbor is any agent, either part of a social network or not, that can be reached within budget on the transportation network (round trip travel). This neighbor list is updated if the transportation network or the travel budget of the agent changes.

Each dyadic association is maintained with a home-based trip; no agents may economize their travels by chaining their trips to maintain social contacts. Enumerating and evaluating all possible trip chains would have been too expensive computationally. If one considers a time step to be the order of a month or a fraction of a year, say a trip to a business meeting, then it is entirely reasonable to consider the visits to be home-based and to occur one at a time. This is probably sufficient for constructing the social network. However, for joint planning models in which several people travel together or meet somewhere, it would be desirable to enable trip-chain calculations, as in Marchal and Nagel (2005).

In each time step, each agent calculates the utility of making a visit to each of his neighbors in a double-buffered calculation: first, each agent plans a list of moves, then the moves (visits to
neighbors) are carried out all at once. The utility is the expected improvement in the agent’s social rank that it would realize by visiting the neighbor, discounted by the travel distance to reach that neighbor. Since the neighbors are all reachable within the travel budget, each move is feasible. But the agent cannot make all the moves, due to the travel budget. The best moves must be chosen. The utilities for the visits are sorted from highest to lowest. A list of planned moves is made for each agent by adding the ID of the target neighbor with the highest utility to the agent’s “move list” and subtracting the travel distance to that neighbor from the travel budget of the agent. This is continued with the next highest utility move, until the “move list” contains enough moves that the travel budget of the agent is exhausted. Once each agent has its move list complete, the moves take place for all agents at once, in a second buffer. In the first time step, there are no social connections yet in the base model runs, and the utility of a visit is undefined for all agents. So the agents are shuffled and the visits are made randomly, within the travel budget. Ties are broken by shuffling.

The social status indicator used is “normalized betweenness centrality” (Freeman 1977). Betweenness centrality is the ratio of the number of shortest paths through the social network which pass through the agent, relative to the total number of shortest paths through the network. It is a measure of how much influence an agent can exert on the accessibility of other agents; thus it is an indicator of this agent’s power in determining the flow of information or goods through the social network. The measure is highly sensitive to whether a link is incoming or outgoing. It is not the same as total degree and this network evolution therefore differs from a preferential attachment algorithm. Adding certain strategic connections to an agent can greatly increase its betweenness centrality. But, in general, any increase in the degree of the agent (the total number of connections) will not decrease its betweenness centrality. Thus, an agent desires to add connections from itself to another agent. In doing so, it bears the cost of maintaining the association; i.e. the transportation trip to meet the desired partner. An agent does not refuse any connection from another agent to itself, as this brings a benefit to the agent which comes without cost.*

The utility of agent $i$ making a connection to another agent is the improvement that can be expected in agent $i$’s social rank: the difference between the normalized betweenness centrality of agent $i$ in the next time step, given a move to agent $j$, and the current normalized betweenness centrality of agent $i$. Betweenness centrality has to be normalized to the network in each time step because it depends on the number of shortest paths in the social network,

* Realistically, the agent would want to attract attachments to itself, as well (incoming links), as in handing out business cards or publishing one’s own professional web page. This could be a very interesting model of a market for social links. But this is not implemented here.
and the size of the network (number of links and linked nodes) changes with each proposed “visit” in this algorithm.

Because a link to an agent has a chance of disappearing if it is not used, the value of updating this link becomes higher the older the link is. That is, the more likely the link is to disappear, the more urgent it might be to re-visit the agent (Table 1). To account for the likelihood of a link disappearing, the future value of a link is discounted in the utility function by multiplying by the probability of its disappearance. The utility expression is, then:

\[
U(visit_{i\to j}) = (\beta_{i\to j}(t+1) \times socialStatus(i | visit_{i\to j}, t+1) - socialStatus(i, t)) \times e^{-\alpha Cost_{i\to j}}
\]

where \(\beta\) is the value of the logistic link removal function at time \(t + 1\) and varies from 0 to 1, \(\alpha\) is the devaluation of the utility with travel impedance, commonly taken as 0.2 (P. Fröhlich, IVT/ETH, personal communication).

Essentially the “social network dynamics” is a generating algorithm similar to preferential attachment (Barabasi 2003) or Watts’ (Watts and Strogatz 1998) small world re-attachment procedure. However, the attachments here are made deterministically with the only probabilistic step being link removal. Another difference is the spatial component: the travel budget only ensures locally maximized social standing (weighed against transportation disutility).

The very simple model is extremely complex because of the dimension of the two networks. The simple parameter space consists of the shape of the logistic link removal probability, the discounting of the travel distance in the utility function, and the travel budget. Other influential parameters include the number, topology and spacing of agent nodes and links, the travel link costs, and the topology of the transportation network.
7. Verification and Results

The base model has a transportation grid network of 100 nodes with link weights equal to 1.0. 36 agents are spaced on this network a distance 2 apart (Table 2), and they have no initial social connections. Agents have a uniform travel budget of 6 per time step. Thus, all agents are able to make one visit per turn to another agent (cost is 2 there, 2 back) and have 2 travel budget left over, which is useless to them. The agents have no socio-demographic characteristics and are identical except for their position on the transportation grid. The agents are spaced evenly, but there is inequality in that the outer agents cannot reach as many others as the middle agents. The outer agents will therefore not likely reach very high social rank scores because they are not as accessible, and a distribution of social ranks is expected, based on spatial considerations alone (Table 3).

Table 2  Road network and links (small blue nodes) and agent locations (large red nodes) illustrates the base case setup. Road links each have cost 1.0.

The graphics in the model (Table 3) are not as good as the hand-drawn figure(Table 2). The social and transportation network in the model graphics occupy two panels next to each other, but they should be considered as mapping onto each other as in the above figure.
Table 3  General disadvantage of a location at the physical edge of society. The model graphics are not yet very nice. The left panel is the transportation network and the right panel is the social network, in which the agents appear in the positions of their home address. The reader’s imagination can overlay the right panel on the left with the key that agent number 35 lives at spatial node 78.

It takes a handful of time steps for the social network to reach an equilibrium in terms of the aggregate social graph statistics. This time interval depends on travel cost and budget. This will be referred to as “spin up”. When the model is spun up, the aggregate social network statistics are stable. A presentation of batch output of this model using different random seeds follows some qualitative discussion of investigations of the parameter space.

The effects of different travel budgets, initial social network topologies, and random number seeds were investigated and the results were qualitatively consistent with expectation. Different random seeds did not change the equilibrium aggregate statistics (see batch runs).

The limited travel budget and the link removal mechanism effectively erase the initial topology of the social network after two link half-lives (10 time steps) such that the initial social network topology does not have an effect in the aggregate. The social network at the last time step and the transportation constraints play a stronger role on overall social network characteristics. But a closer investigation of agent scores (betweenness centrality), which has not been done, might show a meaningful effect of the initial social network topology, even after many time steps. It is clear in the algorithm that the microscopic characteristics of the graph are path dependent. Further research into these results is necessary.
The effect of decreasing the budget below what is needed to sustain relationships is clear: no relationships are established. Higher travel budgets have the same effect as lower (homogeneous) travel costs in the transportation network, and result in a wide distribution of social link lengths (=distance on the transportation network between socially linked agents), meaning that agents are effectively trading off travel for social status, as intended in the model specification. This is an indication that the parameters have been chosen appropriately so that the model is not overdriven by either social rank or distance costs to remain forever in a single-point solution.

The model was run in the base version in batch with 20 random seed values for 30 time steps. It was also run 20 times with the same parameters in a configuration which introduced a set of diagonal links in the transportation network at the beginning of time step 15. The diagonal links have a cost of 1.0 and serve to short-circuit the circuitous routes that the agents have to make on an orthogonal grid. To economize in reporting the results here, the results from both runs are presented together in the graphics. This comparison helps to understand how the model functions, especially how the social network depends on the transportation network. Otherwise the meaning of a single set of conditions has no context.

Table 4  Road network improvement introduced at time = 15: diagonal links also have cost 1.0.

Output from the transportation and social network graphs is written after each time step to three files: an edge file, an agent file, and a graph file containing aggregate statistics. A suite of statistical measures was established to evaluate the state of the social and transportation networks in time and to be able to compare them.
Accessibility: this is a standard aggregate measure in transportation for how easy it is to get to a point on the network. It is essentially the Logit utility expression for travel time to the destination, weighted by the number of trips made there. The accessibility depends on the transportation network and on the attractiveness of the region, as well as the willingness and ability of others to travel to the region. An aggregate measure of accessibility for all agents can be defined as:

\[ A(t) = \sum_i \sum_j e^{-\alpha \text{Cost}_{ij}} \]

with \( \alpha \) and Cost defined as above. Essentially it is the inverse exponential sum of the distances that all agents travel to maintain their social networks, modified by \( \alpha \). Improving the transportation system, placing more agents more densely together, or raising the travel budget would increase accessibility.

Average Path Length: this is the average length of the shortest paths between an agent and all the other agents in the social network. It is an average measure of how close the agents are to all the others in the network, in contrast to diameter (see below). It is the mean value of the JUNG method GraphStatistics.averageDistances. A smaller value indicates either a small number of agents, a high density of connections between agents, or a world in which there are effective shortcuts through the clusters of agents.

Clustering Coefficient: is a measure of how many of an agent’s social contacts are also connected with each other, relative to how many connections are possible between them. A small clustering coefficient is indicative of a centralized star network, and higher clustering coefficients indicate more spread out contacts between agents, a sort of “democratic” or “flat” organization. This is also the mean value of the JUNG method GraphStatistics.clusteringCoefficients, for which the social network had to be temporarily converted to an undirected graph.

Diameter: is the longest of the shortest paths between a pair of agents on the social graph. A small diameter indicates that agents are very tightly connected, similar to a high density of links. A larger diameter means that agents are separated by many intermediaries and that they don’t know each other directly. A high diameter might be expected if the travel budget or the transportation network prohibited agents from knowing many other agents.

Link Length: The round trip travel impedance (distance) between two agents on the transportation network.

Score: The normalized betweenness centrality of an agent or of all agents in the social network at time \( t \).
Table 5  
Accessibility Measure: No build and build with standard deviation over the ensemble of 30 runs

<table>
<thead>
<tr>
<th>Time Step</th>
<th>AvgAccessBuild15</th>
<th>AvgAccessNoBuild</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>10</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>15</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>20</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>25</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>30</td>
<td>110</td>
<td>110</td>
</tr>
</tbody>
</table>

After time step 7, the model is spun up and the accessibility reaches a stable average value with a coefficient of variation of 5.7% due to the different random seeds. The random seeds cause different social connections to form and therefore a different travel demand pattern (origins and destinations) among the agents, and a different sample of travel costs, resulting in the variation in accessibility. As expected, accessibility improves with the addition of the diagonal road, which shortens the distance travelled between agents located close to the new road. The change is 18% and the resulting accessibility is even more stable, with a coefficient of variation of only 4.4% due to random variation. The reduced response to random changes in the social network indicates that the low-cost road network is a strong influence on the formation of social links.

The average path length and clustering coefficient show a similar pattern of initial spin-up and equilibrium. The average path length of the base case is 3.18(0.07) and with the road improvement it is slightly lower at 2.98(0.07), which is a statistically significant difference. Agents can reach each other on shorter paths in social space due to the new road. The
The global clustering coefficient versus path length shows an unexpected trend, however: clustering decreases with increasing path length. Watts (1999) work on small world networks would have predicted a monotonically increasing logarithmic-shaped curve. As friends of friends become acquainted, and clustering decreases, the path length through the social
network decreases toward that of a random graph. The two sample points in Table 6 indicate that this model does not follow the same trend as Watts’ model. In this model, the path length increases with lower clustering. This is likely due to the directed links used in this model causing longer path lengths. The role that the link removal and the travel budget play in causing this result also need to be looked into.

Table 7 Clustering Coefficient versus Average Path Length

<table>
<thead>
<tr>
<th>Build</th>
<th>No Build</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.25</td>
<td>2.50</td>
</tr>
<tr>
<td>2.75</td>
<td>3.00</td>
</tr>
<tr>
<td>3.25</td>
<td>3.50</td>
</tr>
<tr>
<td>3.75</td>
<td>4.00</td>
</tr>
<tr>
<td>4.25</td>
<td>4.50</td>
</tr>
</tbody>
</table>

The mean diameter of the social graph is 7.13(0.3) with no road improvement, and 7.09(0.4) with it, which are not statistically different. Thus, the maximum shortest path length in the social network is not affected by the changes to the transportation network.

A histogram of degree distribution (in degree plus out degree) shows that the change to the transportation network caused a few agents to become extremely highly connected while the rest of the distribution remains stable. This is another indicator of clustering and is likely a measure of the range of the effect of the new road section: were the effect far-ranging, all agents would have more connections. The fact that only a few agents have higher connectivity indicates that the road improvement’s effects are not able to propagate through the social network. The reasons for this should be investigated at the agent level.
What is interesting, however, is that social clustering of identical agents can be induced by making certain agents easier to access physically. This is of course due to the strong assumptions about maximizing centrality used in the utility function, but it is also an indication that social science studies including degree distribution might need to incorporate elements of spatial search in addition to social concepts like homophily before arriving at conclusions about the causes of clustering.

Table 8  Histogram of degree distribution at time 29 without (left) and with road improvement (right)

Table 9 illustrates the distribution of the value derived from travelling: the ratio of the social status score per unit of travel for all the brand new links that were formed in the runs. The gold-colored bar graph with the road improvement is positively-skewed and shows relatively more agents at both extremes and fewer in the middle values. This shows that a few agents derive a high ratio of social benefit to travel distance by the addition of the new road, whereas a high number derive less valuable travel rewards than they had enjoyed without the construction.
When no road is built, the ratio of score to travel (blue bars) is symmetrically grouped around an average value. The average score/unit of travel is 19.1(9.4) without the road and 18.6(16.2) with it. This shows that, although the mean values of score per unit distance cannot be differentiated with statistical significance, the standard deviations are significantly different and indicate a highly uneven distribution of high- and low-value social links in the “build” case relative to the “no-build”.

Another question asks how the travel distance is distributed. The link lengths are distributed differently with the road build because agents have a shortcut which allows them to split their travel budget on two routes. They have a choice to travel 2, 4, or 6, whereas in the base model, they can only travel 4 because no other agents are reachable at distance 2 or 6. Table 10 shows the number of social links versus their length on the transportation network that were made over the last 10 time steps (equilibrium) in 20 ensemble runs.
The length is an integer value because the travel costs on the transportation network are integer values in this example. When the shortcut is built, certain agents find it advantageous to travel the full 6 to a more valuable agent. The neighborhood within distance 4 clearly was not utility maximizing for these agents (by construction). Almost twice as many trips of distance 2 are made as distance 6, and more trips are made altogether if transportation is improved. The extra gains in utility are made possible by enabling agents to use the residual travel budget or to travel to agents who are farther away.

One last statistical cut through the model asks how agents distribute their visits. Table 11 shows the number of unique agent pairs versus the probability of a visit taking place between that pair on each turn. Only the last 10 turns of the base case are shown (equilibrium case).
Directed visits were made between 210 agent pairs over the last 10 time steps, resulting in 13,000 visits in the 20 run ensemble. The plot shows the number of unique agent pairs versus the probability of a visit each turn. For example, at the mode of this distribution, for 52 agent pairs, there is a 35% chance that a visit will occur each turn. These are relationships of fairly high value relative to other alternatives available to the agent. Renewing a relationship after only $1/0.35 \approx 3$ turns means that there is only a 20% chance of the link disappearing if it is not renewed (see Table 1). With the devaluation of utility according to the expected likelihood of link removal, the marginal utility for renewing the connection after only 3 turns is 20% of the actual value of the new link. The other available relationships are therefore worth less than 20% as much as the existing relationship for these pairs.

At the high end, for 2 agent pairs, it is almost certain that a visit occurs each turn. These relationships are very high valued either because they bring very high utility or because they are the only relationships available within the travel budget.
8. Implications and hypothesis tests

Clearly there is a relationship between the spatial arrangement of the homogeneous agents’ home locations and the emergent social networks, given a transportation budget, a behavior rule, and a transportation network, all of which are homogenous. A large degree of inhomogeneity emerges however, because of the spatial boundary conditions. In general, agents are better positioned to capitalize on social network flows if they can reach more agents, namely, they live in a position near the center of the transportation network. Agents at the edge cannot participate as much in the flow of information.

The distribution of trips to maintain social status is complicated to describe. Although the degree distribution at equilibrium with the homogenous travel costs looks Normal and is not a power law, certain links have much higher value than other links, and are visited more frequently than others, indicating a kind of lock-in to a cluster, or at least a dyad, which may be related to parameters which were not included in a sweep in the batch runs: budget, number of agents vs. density of transportation network, or the old problem of spatial boundary conditions, etc.

The social network can be manipulated solely by changing the transportation network. The changes occur in the expected directions except perhaps for clustering, where longer spatial relationships were expected with decreased travel costs. Increasing the accessibility asymmetrically increases the total distance travelled, increases the number of social visits, raises the dispersion of benefits, but doesn’t change the net benefits in the social system. The data written out at agent and edge level can be used to analyze the impacts, but it is very complicated and has not been used extensively, yet.

The hypotheses can be addressed in part:

1. The size of a person's "contact space": The travel budget limits the maximum number of contacts. A calculation of the number of visits to each neighbor divided by the number of reachable neighbors was not done, but could be with this model. This investigation would become more interesting if agents could get to know each other through their social networks rather than by travel alone, a kind of substitution effect of social networks for travel.

2. People maintain contacts to more non- or partially overlapping networks (clusters): Bipartite social networks are not modelled here. But clustering does decrease with increased accessibility. Information can be exchanged more democratically, spread more evenly.
3. Contacts are more selective; social contacts are more valuable: Certainly an agent jumps to a more valuable agent as soon as the budget allows, and neglects local contacts unless they become more valuable. The average social power fell without statistical significance when the road improvement was added. More importantly, the standard deviation of scores showed extreme differences after the transportation improvement. This higher standard deviation across agent choices could indicate higher selectivity of social associations. Without heterogeneous agents, more cannot be said about selectivity in choice of social associations.

4. Weak ties, visit frequency: Without the ability to make friends via the social network (no travel cost), the agents in this model remain clustered in spatial cells, like Watts’ caveman model. How to discover and/or enable weak ties in this model remains to be addressed. A before/after study was not made on the frequency of visits, but analysis of the link length distribution shows that shorter-than-average trips are twice as likely as longer-than-average trips if the road shortcut is added. This shows that the total number of trips increases rather than becoming more efficient and decreasing.

5. The number of social activities an actor is involved in increases with his social capital: A plot of agent score versus degree or likelihood of visitation was not made. It would be feasible to look at this question with this program.
9. Next Steps

Now that the basic framework of a linked social and transportation network is running, and a basic graphical interface exists, with statistical output, it is desirable to simplify the basic model function and carefully analyze its characteristics, to place this model in the context of existing network generation models. The small-world generation models are based on a rewiring probability. This could be easily reproduced here by replacing the deterministic utility with a random utility model that chooses the new link to be established based on travel utility, without regard to a transportation budget. The betweenness centrality term in the utility could even be replaced by a simple degree centrality, bringing the model in line with a preferential attachment model. Setting $\alpha=0$ should recover a scale-free small world. Once the output is examined in a similar way to other published network growth models, steps can be retraced to evaluate the effects of complexities like travel disutility, a travel budget, spatial boundary conditions, link removal conditions, initial social network topologies, and other measures of social status like betweenness centrality.

The underlying transportation network is about modelling costs of links. Certainly the typical phenomena modelled in network growth models: worm ganglia, cell respiration (Newman 2003), etc., do not occur costlessly. These models could be enriched by a cost function. The form of this model is particularly interesting for such applications because it enables dynamic coupling of two networks.

Further application in transportation will require a more realistic route-finding algorithm than a Dijkstra shortest path. Of particular interest would be the ability to calculate intermediate meeting points for agents, such as a public park or a restaurant. This might require a kind of negotiation game between the agents, based on travel impedances, social standing, etc. It would be a complement to the model of Marchal and Nagel. Once an agent model with this degree of refinement is running, it becomes realistic to begin cost/benefit analyses of adding such treatments to larger scale transportation models.
10. References


Brandes, Ulrik http://graphml.graphdrawing.org/


