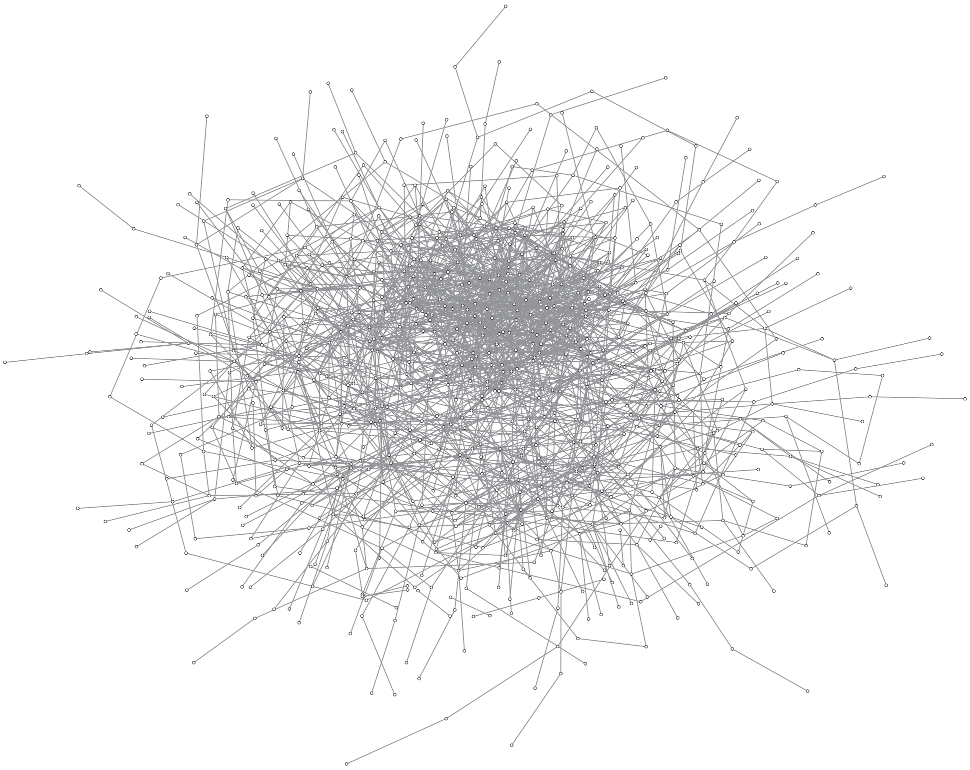


Social networks in transportation microsimulations

Jeremy Hackney, IVT/ETH

Conference paper STRC 2007

STRC

7th Swiss Transport Research Conference
Monte Verità / Ascona, September 12. – 14. 2007

Title of paper

Jeremy Hackney
Institute for Transport Planning and
Systems/ETH
Zürich

Phone: +41(0)446333325
Fax: +41(0)446331057
email: hackney@ivt.baug.ethz.ch

September 2007

Abstract

The term "social networks" has attracted attention in many disciplines and in the popular media since the publication of Watts' and Strogatz paper on "small worlds" in 1998. Nodes of the social network represent entitites and weighted edges represent relationships and their strengths, explicitly accounting for interdependencies. Social networks are rarely studied in a geographic context, and, while transportation researchers have long recognized that travellers' decisions are interdependent, social networks are rare in our field. First, the common statistical tools require the assumption of independent decision makers. Second, real maps of social, work, and business interactions have not been made at the scope required for transportation studies. Agent simulation is one method available to construct the interdependencies between travellers and activity-travel choices. This research uses the basic tools of transportation science: utility maximization, activity plans, and generalized travel cost, to construct social networks for a geographically distributed population of agents. This network could then be used in a further step to modify travel demand (location, activity, and/or timing). This paper summarizes the use of social networks in the field of transportation planning, illustrates some of the issues of investigating complex socio-spatial dynamics, and presents results of a simple simulation. Patterns seen in real social networks can be reproduced with reasonable behavioral assumptions. The paper concludes by discussing the computational challenge of scaling the models to large samples of agents, and the integration of the socializing dynamic with a large-scale microsimulation.

Keywords

behavior microsimulation – social networks – activity-based planning – travel behavior

1. Introduction

Travel utilities (and decisions) are almost always modelled as being individually independent across decision makers: a function of the traveller's own characteristics and the attributes of modes, routes, and destinations. The fact that people plan trips and activities jointly, depending on the trips and activities of other people, has been ignored with few exceptions (Kurani and Kitamura 1996). This compromise has been reached consciously and with good reason, including convenient econometric model estimation tools (e.g. Ben Akiva and Lehrman 1985) and the lack of datasets detailing decision-maker interactions (Manski 2000). Representing behavior as being entirely independent succeeds as long as it considers regulary occurring, relatively inelastic equilibrium behavior, like work (non-discretionary) trips, where the boundary conditions of the problem dominate the solution. Travel preferences in this case can be assumed to be explained by the sociodemographic characteristics of the traveller and by generalized travel costs; any deeper processes of orchestrating the trip take either a long-term character exogenous to the model (Hensher 2002), i.e. choice of home/work location, or they play an insignificant role due to the relatively limited flexibility of the traveller. More flexible (discretionary) behaviors like leisure, or partially flexible activities like shopping, are explained by adding variables like taste heterogeneity and cohort or habitual behavior effects, either as random coefficients or based on panel datasets or assumed distributions.

While such extensions improve the statistical specification and the conceptual validity of the models somewhat, they neglect accounting for agent interactions. This rules them out for studying important processes of real behavior, and may limit their usefulness in prediction, as well. The model output may be useful, but it may be for the wrong reason. It has been observed that peer, competitor, or group influence on decisions can be a strong component in consumer behaviors from technology adoption to restaurant choice, and that these so-called 'network effects' are even the primary driver of some markets (Shy 2001). Analytical models for these behaviors are well-described. Analogously, social processes like information exchange, imitation, and simply enjoying the company of others could be playing a role in travel decision-making precisely in the areas of (what appear to be) flexible choices and spatial learning (choice set formation) that are now either poorly explained or which determine computational limitations (e.g. destination or route choice set size) in the transporation field. Further insights are not possible in current frameworks without taking the plunge to incorporate agent interactions.

Endogenizing interactions amounts to explicit accounting for who influences whom, when and how much. This is a dimension of disaggregation (for each of *n* observations one might conceive of up to *n-1* social connections) which opens the possibility of resolving spatio-temporal behaviors that would otherwise be subsumed by sociodemographic explanations.

But there is a major drawback to methods using social interactions, because it becomes necessary to know the specific social connections (social network), as well as the direction and strength of influence on an individual's decision making that is communicated by this network: the so-called *reflection problem* (Manski 1993; Bramoullé et al. 2007). This phrase refers to the difficulty an observer has in determing whether a person moves his image in a mirror or vice versa: without knowing the relationship between the two people in view, i.e. that the mirror exists, the observer cannot separate cause and effect. The metaphor applies to the cognate sociological problem of determining to what extent the correlated decisions between an individual and a group is a result of a process of choosing peers with the same attitudes, versus a process by which members of a group influence one another to conform to some group norm. In order to get the right answer about how the society functions, precise knowledge about, or hypotheses of, social topologies and the dynamic processes at work in the system are necessary.

Both communication and person-to-person encounters are crucial for the maintenance of social networks and the social capital that they enable (Larsen et al. 2006). A reciprocal interaction between travel/communication and social networks is therefore a logical starting postulate: social networks generate communication/travel, but communication/travel opportunities enable the spatial spread of the social networks. But these are rarely observed (Manski 2000).

1.1 Goals

This work accepts that observations of social interactions in the transportation context are sparse, and employs a micro-simulation approach to achieve basic understanding of a realistic range of interactions between travellers that could potentially be influencing activity-travel choices. The leading research question is how to generate and characterize social connections (networks) and activity spaces concurrently such that they are consistent with both social theory and travel behavior theory. The micro-simulation is a modular structure based on the MATSim Toolbox (summarized recently in Rieser et al. 2007) and it enables the researcher to conduct and document experiments by inserting desired algorithms and/or datasets for the three main elements of study: socializing, geography, and travel behavior, and to adjust the strength of their coupling. Using multi-agent simulations to model inter-actor relationships that lead to macro-scale system properties is a combination of deductive and inductive methods, sometimes called “generative science” (Sawyer 2004). Agent modeling allows researchers to control and experiment with microscopic behavior and observe the emergent macroscopic system (Bankes 1993; Axtell 2000). There are virtually no limits to the assumptions a researcher can make in the module, though the success of algorithms will depend on computing limitations (and the justification for the assumptions).

This paper describes the implementation of the social network module within the MATSim-T micro-simulation framework. A summary of what is known about social networks and the geography of social travel is next, followed by other work on social networks in transportation. Then the MATSim-T project is summarized, and the social network module is described. Sample results illustrate some of the possible configurations of the module. Future work will include application to large synthetic populations on a realistic road network, after which the social network statistics and the activity space dimensions will be compared with what is known from the data to elucidate relationships to use in future work in surveying or policy analysis.

2. What is known about travel and social networks

2.1 Real social networks

A social network (graph) is a mathematical expression referring to a set of nodes (vertices), representing people, and links (edges) representing well-defined relatinships between the people. The links can be valued and directed (arcs) to represent relationship strengths, and these values can change in time. Network statistics are counting procedures focused on certain constellations of links and nodes: for instance clustering (essentially a normalized count of triangles), shortest path length (minimum number of links connecting two nodes), and degree (number of links entering or exiting a node).

The evidence from samples of real social networks indicates that they are neither well-represented by models in which relationships are equally likely between all individuals, independent of physical distance (Erdös/Renyi network), nor by models of lattices, in which people only know their neighbors. Instead, real social networks fall somewhere in between these two well-studied theoretical structures, presenting a so-called "small world" structure with properties of both types of networks, in which a few individuals belong to local social clusters, but also to other distant clusters, thus providing shortcuts through society (Watts 1999; Newman et al. 2002). Small world network statistics are characterized by a higher clustering coefficient than an Erdös/Renyi graph, and a small average shortest path length (log (number of nodes)).

In addition to small world properties, observed social networks exhibit a distribution of degree which follows an exponential distribution or a distribution between exponential and scale-free (Dorogovtsev and Mendes 2003), meaning that there is a representative average number of relationships per person, but also a tendency for certain individuals to attract social connections at much higher rates than other individuals. People preferentially associate with certain other people, for whatever reason (Barabasi et al. 1999), and the number of friends is correlated with the number of friends that friends have . This is another major difference from Erdös/Renyi networks, which have a symmetric Poisson degree distribution and no correlations.

Social networks used in/emerging from transportation studies should exhibit these realistic characteristics, or else network phenomena will not have the proper heterogeneity due to clusters, or spread quickly enough, due to short paths. The networks will also not be susceptible or resilient enough to failures that can occur to nodes or links. In short, many social networks may be considered "realistic", but the answer will be wrong if Erdös/Renyi networks are used.

2.2 Datasets of social travel

A survey of the datasets available to study the geography of social interactions is difficult due to the small number of studies and the wide range of literatures that has to be searched. The microsimulation is to some extent an alternative to large datasets, but even so, data is needed for initializing social networks, for descriptions of travel-related interactions, and for comparison with measures of social networks and activity behavior with simulation output. Ideally, the data would include the typical social network measures like degree distributions and clustering coefficients, but also behavioral measures like the frequency and location of face-to-face meetings, the activity types associated with these meetings, the number of participants, their relationship, the planning process, and so on.

(Axhausen and Frei 2007) summarize a wide range of studies about the geographical distribution of social contacts. Universally, face-to-face contact frequency falls with the inverse of distance. However, microscopic rules are more difficult to extract. Datasets focusing on the social influences on travel behavior ((Silvis et al. 2006), (Axhausen and Frei 2007), (Mok et al. 2007), (Carrasco and Miller 2006)) are characterized by detailed questions about behavior posed to small numbers of travellers. They are useful for indications but not for statistical modeling or to provide strong support for hypotheses. In epidemiology, (Christakis and Fowler 2007) constructed a geocoded dynamic social network of spatial interactions of 5124 people over 35 years and analyzed the role of network statistics in the spread of obesity. A geographical analysis is pending. (Rothenberg et al. 2005) analyze the distances between sex partners in an HIV study.

These data, as they can be acquired, will be used as possible to specify microscopic behavior models and compared with the generated networks at a later stage in the project.

2.3 Application of social networks in transportation

2.3.1 Econometric studies

Statistical analyses of social network effects published in the transportation literature follow the theory of (Manski 1993) and derivations of (Durlauf and Cohen-Cole 2004). (Dugundji and Gulyas 2003; Páez and Scott 2004; Dugundji and Walker 2005) generated social networks on the basis of a number of factors (common zip code, common residential or work zone, common workplace, sociodemographic categories) and used them to estimate econometric discrete choice models estimated on revealed preference data. The studies show that the endogenous normative opinion of a peer group in certain social networks can have significant explanatory power for mode choice and trip generation. It is important to note, however, that neither microscopic behavior nor individual interactions could be taken into account in these models.

2.3.2 Microsimulations

Microsimulations of social networks in the transportation literature include work from (Arentze and Timmermans 2006), (Hackney and Axhausen 2006), and (Marchal and Nagel 2006). (Arentze and Timmermans 2006) present a fully developed concept for social interactions and activity patterns based on the ego-centric (personal) network, including abstractions of homophily (McPherson et al. 2001), social need, and satisfaction. Their utility functions maximize the value of ego networks within the total discretionary time budget of the agent. Tests were limited to small numbers of agents, so there are no summary statistics of the social networks, and it is noted that the results are very complex.

(Hackney and Axhausen 2006) report similar problems analyzing the complex results from a model of socializing in which travel cost is weighed against participating in social activities in the utility function, and agents exchange information with their affiliates about where other socializing opportunities exist. Established friendships and familiar activity spaces are favored over random encounters, spatial discovery, and staying home, with visits to places recommended by friends (where friends of those friends are likely to be) having intermediate utility. The model is a rudimentary activity travel model with homogenous agents and no explicit valuation of a social network, intended as a template for realistic models with estimable utility functions. Despite its simple logic, the geographic provenience of the agents is a step up in complexity from existing network generation algorithms. Statistical analysis of the networks indicates exponential degree distributions and probabilities of affiliation proportaional to the inverse of separation between alters' home adresses. This model's algorithms are not scalable to models larger than several thousand agents.

(Marchal and Nagel 2006) have modeled the spread of information about secondary location choice (shopping) along the affiliation network of co-located coworkers to accelerate the learning curve of agents. However this work used Erdös/Renyi social networks and served above all to illustrate the feasibility of the approach in a large-scale microsimulation.

3. MATSim-T

MATSim-T (Multi-Agent Transportation Simulation Toolbox, see (Rieser et al. 2007)) is an activity-based demand generation and dynamic assignment simulation. Behavioral modules allow the researcher (or planner) to simulate agent learning and adaptation by letting agents follow various strategies to maximize the utility of activity plans and relax the system. The program is Object Oriented Java under public GNU license, and is downloadable at www.sourceforge.org.

3.1 Program structure

The program is structured in layers (Balmer 2007) of arbitrary resolution that point to each other: for example, a hectare grid, traffic planning zones, municipal boundaries, and a road network. Aggregations up and disaggregations down the layers are performed automatically. In setup, facilities in which to perform activities are placed on the appropriate layer and aggregated accordingly to other layers, and especially to the network layer, where calculation (assignment) takes place. The exact details are more complicated and may change soon, but this description suffices to understand the simulation. A synthetic population is assigned home locations and a temporal plan consisting of a series of activities to be accomplished, and an initial schedule for the activities. The agents are also given knowledge about some portion of the facilities in their world, akin to an activity space. Finally, the locations of an agent's activities (factilities) are allocated from the agent's knowledge.

3.2 Assignment and replanning strategies

The agents initially negotiate what would be the shortest path on the non-loaded transportation network between their activities and experience large traffic delays. An iterative replanning procedure lets the agents adjust their plans in any number of dimensions in order to reduce the amount of time they spend outside of participating in their activities. The method employed is called, "partial relaxation" in which only a portion of agents may replan their day, while the rest stay with their current plan. It is used to avoid the situation in which all agents would jump to the same uncongested paths at once, each iteration, which causes a toggling or flip-flop effect in the assignment. The partial relaxation reaches the same goal eventually. As it is implemented, it is possible to let portions of the population use different strategies for optimizing their plan. Thus the researcher may, if desired, perform a study of a kind of evolutionary strategy for optimizing activity plans. Strategies may also be used for a certain number of iterations and then switched off or over to another strategy (usually an optimization trick). The strategies available in the standard MATSIM-T package include re-routing, altering the departure time from activities, and altering the duration of activities. These, in turn, can be performed using either random mutations, genetic algorithms that combine plans, or recently, landmark navigation.

3.3 Plan scoring and iteration

The current utility function awards an agent positive utility with diminishing returns for performing an activity, and linear negative utility for travelling. It also penalizes linearly for arriving too late or too early at an activity and for performing the activity for too short a time. When a plan is altered by the strategy, its new anticipated utility is calculated. A certain number of plans and their scores are kept in memory for each agent. The plan that is carried out by the agent next iteration is determined by a selector function. The researcher can choose from a number of selection functions to use for each strategy (strategies are paired with selector functions, but all use the same utility function).

3.4 Time horizon

The temporal horizon for plans and thus simulations is not limited, though a 24-hour day is usually run in the tests. That is, the activities are performed within the time horizon, and utility is reset again when the plan is run through the replanning module each iteration.

4. Social Networks Module

4.1 Overall

Introducing social networks into MATSim is an application as well as an extension of the toolbox. The agents in MATSim interact only in traffic, but not to plan or optimize their activities. The first modification is to attach an Ego Network to each agent. The Ego Network is simply a list of friends and a list of the links to the friends. The Social Network consists of a list of pointers to the links between agents. This way, the researcher can access links and people using either piece of information. Finally, the Social Network and the Ego Network are inserted into the agent's Knowledge (previously only geographic knowledge, now social knowledge, as well).

Because of the dynamic interaction between geography and social networks, they are inserted into the replanning section of MATSim. Taking full account of traffic congestion is not necessary for generation of social networks, so this structure may be unnecessarily costly (dynamic assignment can be replaced with a shortest-path search). However, it will be clear in a moment that moving the social network-influenced plans to the "initial plans" stage of MATSim does not necessarily save time.

4.2 General social network dynamics used

The geographic social network dynamics used here consist of five basic steps:

1. Initial social connections (write out net with statistics)
2. Spatial (face to face) encounters
3. Modification of social connections as a result of encounters and calculation of network statistics (write out net with statistics)
4. Other information exchange on social network, not requiring face-to-face encounters
5. Modification of plan

These steps are to be repeated until a desired stopping point, preferably in equilibrium (i.e. the number of links == average degree is constant).

4.2.1 Initialization

The initial social connections may take any topology the researcher desires to generate. At this stage, the researcher must also decide whether networks are directed or not (are acquaintances automatically reciprocal or can information flow asymmetrically). The initial network is given a name that is used as a keyword in the configuration file to load the appropriate Java class where the generation algorithm is to be written. At the moment, only "random" (=Erdös/Renyi) or "empty" have been programmed, of the hundreds of plausible statistical and microbehavioral algorithms that have been published. This initial network might instead conceivably be based on household relationships, extra-household family, work colleagues, or other index of affinity. Note that the selection of characteristics for a Person in MATSim is small (here an excerpt from the .XML file written by MATSim: <person id="1" sex="f" age="20" license="no" car_avail="never" employed="yes"), and to use social network generation algorithms on detailed sociodemographic information will require extending this data structure and, in addition, working directly with the Census data instead of the MATSim synthetic population.

4.2.2 Spatial encounters (face-to-face)

The spatial encounters are based on the plans of the agents. They may meet only when they visit the same facility. A time window is not yet programmed, but this may be easily added to the algorithm because activities have start and end times. (Miller 2005), for example, contributes a new comprehensive view on time window constraints on behavior. At the moment, when agents meet, their social contact is strengthened if they know each other already. If not, the only other interaction that has been programmed is "random" befriending with a saturation effect. That is, agents befriend one another at a random rate, depending on the activity type (both of these are determined by parameters set by the researcher in the configuration file), with decreasing probability depending on how many friends they have. At the moment, there is automatic mutual agreement to befriend. However, the agents may interact in any other way and with any other criterion that the researcher wants them to during their interaction (e.g. pass information, introduce friends), as in the network initialization, by writing the appropriate lines of Java.

4.2.3 Modification of the social network

The strengthening of bonds and adding of new bonds has been mentioned. But links can also disappear. At present, the keywords "none" or "random" in the configuration file govern this adjustment. "None" means links are not removed. "Random" removal means that social links that that have not been renewed by face-to-face contact may be removed after a certain time threshold (also a parameter, where time is measured in iterations). This could easily be replaced by a probability that grows with time, with degree (as in (Jin et al. 2001)), or a function which incorporates the type of relationship (Burt 2000), but note that relationship type has not yet been implemented in the Link class). The JUNG Java library is used at this point to calculate network statistics, which are written out at this stage for the iteration.

4.2.4 Non-spatial information exchange and unobserved social behavior

Passing information along the social network nonspatially is a method written to account for the fact that not all social activities can be observed in a single day (or week) plan, which is an important shortcoming as far as we are generating these social networks, and the fact that not all social contact requires face-to-face travel, which is more a concern for applying the social network to replanning activities. In this method, agents exchange information from their Knowledge about geography and their Ego Networks about each other. The amount and type of information that is exchanged is controlled by parameters in the configuration file. Currently, only a "random" exchange of information is possible. It is assumed that the introduction of friends to each other occurs when information about other agents is transmitted: i.e. A telling B about C amounts to A introducing B and C. This closing of triads is an important function in social networks that leads to correlated degrees, clustering, small worlds, and exponential degree distributions (Jin et al. 2001; Hackney and Axhausen 2006). If friend introductions took place only spatially, an insufficient number of friends would be introduced to other friends, unless one allowed agents to "invite" one another to social engagements. This non-spatial social interaction is a shortcut to programming this behavior explicitly. Again, one can imagine making the exchange of information depend on a criterion, such as agents exchanging information that they personally find valuable, or that they believe their friend would find valuable (Altenhoff 2003).

4.2.5 Plan modification and scoring

The modification of the plan takes place as described in the replanning step. The number of iterations of social interaction that take place before the agent replans its activity chain can be set by the researcher. One may think of this as building tension in the system as agents collect information about higher-utility alternatives, and releasing it by running the dynamic assignment and replanning the activities. The researcher can experiment with this release valve.

The socialized agents can still use the existing strategy modules for replanning, but these do not take advantage of the social network. A special "strategy" module has been written to use the social network by randomly modifying the location of activities within the plan with locations from the agent's knowledge. The knowledge-based replanning algorithm replaces the location of an activity with a certain activity-specific probability, set by the researcher in the configuration file. For instance, it is very unlikely that a person changes home location, slightly likely that he changes work or school location, but highly likely that he changes leisure or shopping locations. Many other replanning strategies come to mind that could be used by the agents, for instance, attempting to negotiate with alters to time their activities together at the same places, or to coordinate similar routes (simulate carpooling).

At the moment, standard utility function is used for individual plan scoring. More is not needed for the location choice replanning module. The feedback of socializing into plans occurs by information exchange about geography across the social network, which changes as agents travel through the landscape. The simplest utility function was chosen initially to keep the analysis of this dynamic simple. It is conceivable that the explicit incorporation of socializing utility would alter the feedback and make interpretation much more difficult. It will be necessary to include social factors in the utility function if other replanning strategies like coordinating schedules are to be studied. However, the implementation of such a utility function is not clear. One must decide, for example, how to weight social factors with respect to each activity. Some leisure opportunities, like the movies, rely on the stimulation of having many people present who are *not known* to the ego (i.e. the sheer number of others present might give positive utility), while other leisure opportunities like a card game might depend on a small number of *known* individuals being present. Experimentation is made mroe complicated by the added parameters. However such questions form the essence of social travel and must be addressed sooner or later.

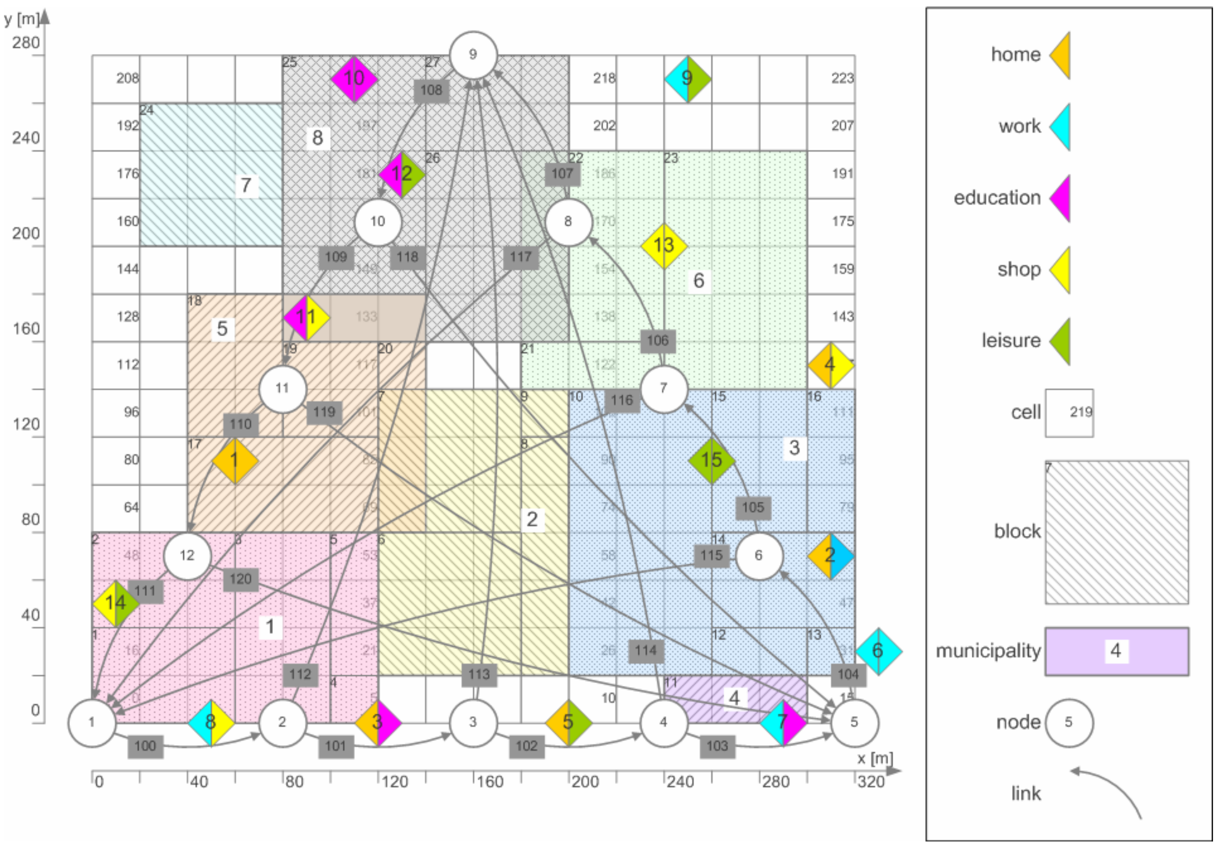
4.3 Output

The variables and format to be output has not been finalized but is in a state for producing quick statistical analyses of runs. All travel information (all movements on the network) is output each iteration of the dynamic assignment, so traces of movements in space can be analyzed. Routines exist in the Toolbox to calculate activity spaces and to plot movements in GoogleEarth. Emphasis has not been placed thus far on spatial results in the test phase of the social network calculations because the test case is not very geographically interesting. Degree is output in text format and the entire social network is output each iteration in a format readable in the Pajek network analysis software package (Batagelj and Mrvar 2003). The output format used earlier by (Hackney and Axhausen 2006) enabled concurrent socio-spatial analyses using postprocessing code written in R, but this program was focused on ensemble analyses of the iterations and has a different structure. Graphical output of simultaneous geographic and social output, written to disk each iteration, will definitely be incorporated.

4.4 Test Scenario

The test scenario uses the TriangleTest world packaged with MATSim-T (Figure 1). 1008 agents are initialized with day plans of up to 4 randomly-sequenced activities each, beginning and ending at the agent's home location. The agents are assigned randomly to a home location, and the locations of their activities are also random. This world contains a rather low number of facilities, which is not ideal for generating geographic social networks, yet it is a very complicated collection of links, jurisdictional layers, and data layers (network and grid), all of which offer an aggregation basis for the analysis of output. This is a foreshadowing of the complicated process of sifting through output of geographic and social dimensions in a realistic scenario.

Figure 1 The geographic world of TriangleTest



Source: (Balmer 2007)

The scenario presented uses a configuration file (Figure 2).

Figure 2 The XML test configuration file, social network module

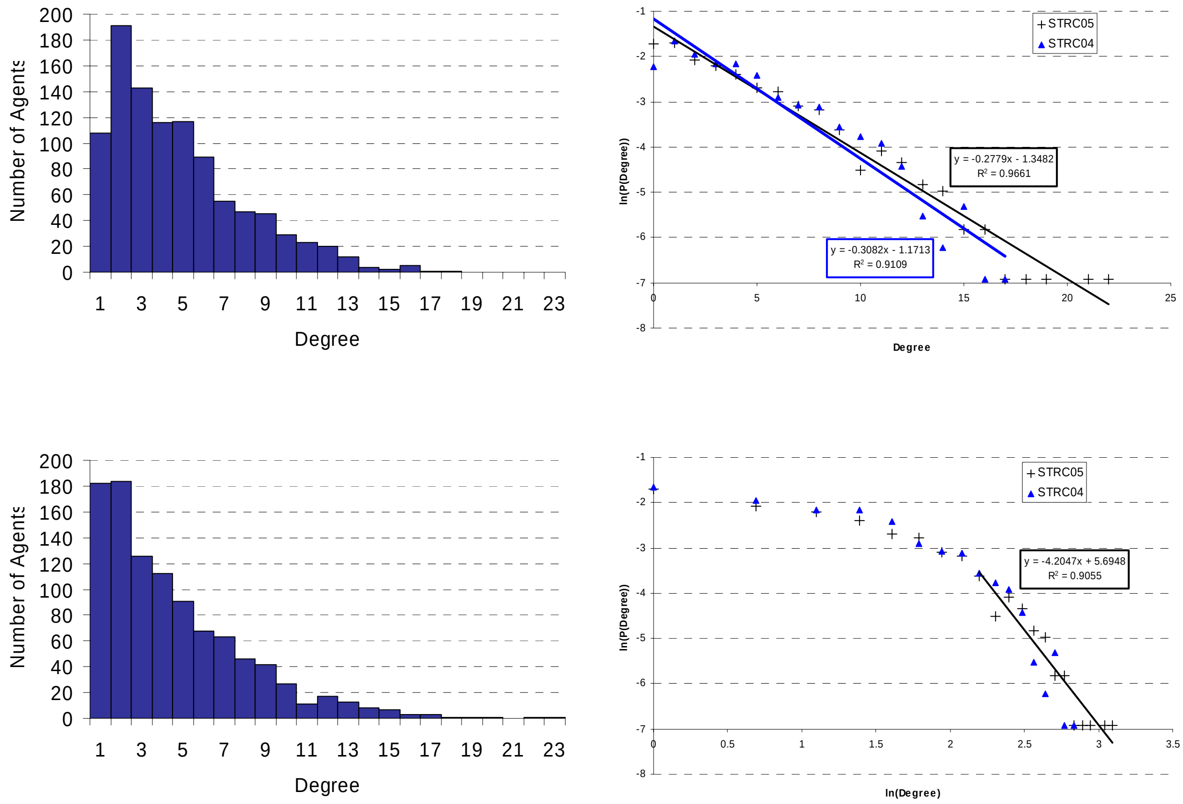
```
<!-- ====================================================================== -->

	<module name="socialnetwork" >
		<param name="degree_saturation_rate" value="0" />
		<param name="edge_type" value="UNDIRECTED" />
		<param name="factype_ns" value="leisure,shop,education,work,person" />
		<param name="fract_ns_interact" value="0." />
		<param name="fract_s_interact" value="1." />
		<param name="kbar" value="0" />
		<param name="max_sn_iter" value="100" />
		<param name="nonspatial_interactor_type" value="random" />
		<param name="num_ns_interactions" value="1" />
		<param name="outputDirSocialNets" value="C:/Documents and Settings/jhackney/My Documents/sandbox00/vsp-cvs/devel/matsim/matsimJ/output/socialnets/" />
		<param name="prob_befriend" value="1." />
		<param name="replanning_interval" value="1000" />
		<param name="s_weights" value=".01,.05,.005,.05,.005" />
		<param name="socnetalgorithm" value="random" />
		<param name="socnetlinkremovalage" value="5" />
		<param name="socnetlinkremovalalgorithm" value="random_link_age1" />
		<param name="socnetlinkremovalp" value="0.05" />
		<param name="socnetlinkstrengthalgorithm" value="constant" />
		<param name="spatial_interactor_type" value="random" />
		<param name="switch_weights" value="0.0,0.01,1.0,0.01,1.0" />
	</module>

<!-- ====================================================================== -->
```

The social network (1008 agents) is initialized with no connections and runs for 100 iterations. The scenarios illustrated in the results differ by changing the variable in ***bold italics*** from 1000 to 1, i.e. never replanning the location of activities versus replanning each iteration of social interaction.

4.5 Results

Aggregate results are shown in Figure 3 of degree distribution for the final iteration of the emergent social network specified in Figure 2. These distributions are output as delimited text files during the run. Other statistics are calculated for the moment in postprocessing. More in-depth results will be calculated and reported hand in hand with the verification process. The process reaches an equilibrium (determined by zero variation in the average degree over a number of iterations) so the results are not illustrating a growing network (which might be expected to have an exponential or scale-free degree distribution).

Figure 3 Analysis of degree distribution for two simulated networks



The histograms on the left show the degree distribution of the 100th iteration for the two runs of Figure 2. The locations of activities were adjusted using knowledge from the social network in the lower run (STRC05) but not in the upper run (STRC04), i.e. in this run, agents met and befriended each other continuously at the same locations. The plot on the upper right shows a log-linear plot comparing the distribution to exponential. The plot on the lower right is a log-log comparison with a scale-free distribution (the linear fit is for STRC05). Both runs yield exponential degree distributions but the run with replanning has a slightly thicker tail. The effect is too slight and the network is too small to show a statistically significant difference.

A summary of basic network statistics shows surprising differences, however (Table 1). The average degree is similar. The run with location replanning (STRC05) has a higher number of components (nearly 20% of the network is not connected: these are likely people staying home all day). The diameters are what one would expect from a random Erdös/Renyi network, but the networks are much more fractured. Fewer than 50 components would be expected of an Erdös/Renyi network of this number of nodes and average degree. When one considers that the diameters are measured in subcomponents of 867 (STRC04) and 798 (STRC05) nodes, these are relatively long distances, i.e. the networks are not as tightly connected as random or small-world networks.

STRC04 with fixed activity plans is twice as clustered as a random network, but STRC05 with location replanning is only 30% more clustered than a random network. The link removal algorithm works together with the probability of making friends to form clusters, despite there being no mechanism activated in these runs for closing triads. However, when agents are free to try more locations more frequently they do not forge clusters at a very high rate because they are apparently switching relationships quickly (and old ones are removed at roughly the same rate). Only an investigation in to the duration of friendships could shed light on what is happening: how link removal works together with changing the locations to affect the social networks must be explained.

Table 1 Statistics of the emergent social networks

Run configuration	Diameter	Number of components	Average clustering ratio relative to Erdös/Renyi random graph	Average Degree
STRC04 no replanning	14	125	2.13	3.77
STRC05 location replanning	12	195	1.31	3.57

4.6 Calculations of Run Time and Memory

The calculation for the test run STRC04 without replanning requires 14 seconds on a 1.8GhZ processor. Running the replanning every social network iteration takes 12 minutes. Replacement of the dynamic assignment with an unloaded network or OD lookup table would be advisable. Larger social networks of realistic proportion (say, 25 acquaintances) will burden the computations roughly with degree squared. An increase of the number of agents to the Zurich region (100x) will require running on a much larger machine. Batch calculations are not yet configured.

5. Future/Continuing work

A scenario with more spatial variation is needed to test the geographic distributions. The scenario is prepared and the runs can be made soon. They will follow an experimental plan such that the model can be verified before any attempt is made at comparison with real distributions. Dynamic contributions to scaling (universal) effects versus parameter-dependent effects must be identified and quantified in order to simplify the models. The plan takes the broad form of Table 2.

Table 2 Framework for the experimental plan for microsimulation of geographic social networks

Social Network	Social dynamics	Utility function	Activity plan
Fixed	Geographic information exchange	Standard	Fixed
Changing	Introducing friends	With socializing factor	Replanning
	Optimizing plan with friends		
	Removing social links		
	Directed links		
	Weighted links		

Verification and identification of relevant parameters and model effects are the first priority. Run time and analysis of output, followed by exchangeable output file formats (a DTD standard for XML) will follow before more interesting types of interactions and perhaps even agent games can be programmed.

6. References

Altenhoff, A. (2003) Location Choice: Modellierung des Location Choice Prozesses in einem Aktivätenbasierten Modell mit kooperierenden Agenten. ETH Semesterarbeit. Zürich**:** 18.

Arentze, T. and H. Timmermans (2006) Social Networks, Social Interactions and Activity-Travel Behavior: A Framework for Micro-Simulation. Transportation Research Board Annual Meeting. Washington, D.C.

Axhausen, K. W. and A. Frei (2007) Contacts in a shrunken world. Arbeitsberichte Verkehrs- und Raumplanung 4xx. Zürich, IVT, ETH Zürich.

Axtell, R. (2000) Why agents? On the varied motivations for agent computing in the social sciences. The Brookings Intitute Center on Social and Economic Dynamics Working Paper Series. Washington, D.C.**:** 23.

Balmer, M. (2007) Travel demand modeling for multi-agent traffic simulations: Algorithms and systems. Dissertation. Zürich, ETH Zürich.

Bankes, S. (1993) Exploratory Modeling for Policy Analysis. *Operations Research* 41: 435-449.

Barabasi, A.-L., R. Albert and H. Jeong (1999) Mean-field theory for scale-free random networks. *Physica A* 272: 173-187.

Batagelj, V. and A. Mrvar (2003) Pajek -- Analysis and Visualization of Large Networks. In *Graph Drawing Software*. M. Jünger and P. Mutzel. Berlin, Springer.

Ben Akiva, M. and S. R. Lehrman (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge

Bramoullé, Y., H. Djebbari and B. Fortin (2007) Identification of Peer Effects through Social Networks. Forschungsinstitut zur Zukunft der Arbeit Discussion Paper Series. Bonn**:** 43.

Burt, R. S. (2000) Decay Functions. *Social Networks* 22: 1-28.

Carrasco, J. A. and E. J. Miller (2006) Exploring the propensity to perform social activities: a social networks approach. *Transportation* 33: 463-480. DOI 10.1007/s11116-006-8074-z

Christakis, N. A. and J. H. Fowler (2007) The Spread of Obesity in a Large Social Network over 32 Years. *The New England Journal of Medicine* 357: 370-379.

Dorogovtsev, S. N. and J. F. F. Mendes (2003) *Evolution of Networks*. Oxford University Press, Oxford

Dugundji, E. and L. Gulyas (2003) Empirical Estimation and Multi-Agent Based Simulation of a Discrete Choice Model with Network Interaction Effects. CUPUM'03 Sendai- 8th International Conference on Computers in Urban Planning and Urban Management. Sendai, Japan.

Dugundji, E. and J. Walker (2005) Discrete Choice with Social and Spatial Network Interdependencies: An Empirical Example Using Mixed GEV Models with Field and "Panel" Effects. *Transportation Research Record* XX:

Durlauf, S. and E. Cohen-Cole (2004) Social Interaction Models. Working Paper, Economics Department, University of Wisconsin. Madison.

Hackney, J. and K. W. Axhausen (2006) An agent model of social network and travel behavior interdependence. 11th International Conference on Travel Behaviour Research. Kyoto.

Hensher, D. A. (2002) Models of Organisational and Agency Choices for Passenger and Freight-Related Travel Choices: Notions of Inter-Activity and Influence. The Australian Key Centre in Transport Management Insititue of Transport Studies Working Paper Series. Sydney**:** 28.

Jin, E., M. Girvan and M. E. J. Newman (2001) Structure of growing social networks. *Physical Review E* 64: 046132-1-o46132-8. DOI: 10.1103/PhysRevE.64.046132

Kurani, K. S. and R. Kitamura (1996) Recent developments in the prospects for modelling household activity schedules. A report prepared for the Los Alamos National Laboratory. Davis, University of California.

Larsen, J., J. Urry and K. W. Axhausen (2006) *Mobilities, Networks, Geographies*. Ashgate, Aldershot

Manski, C. F. (1993) Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies* 60: 531-542.

Manski, C. F. (2000) Economic Analysis of Social Interactions. *The Journal of Economic Perspectives* 14: 115-136.

Marchal, F. and K. Nagel (2006) Computation of location choice of secondary activities in transportation models with cooperative agents. *Transportation Research Record* XX:

McPherson, M., L. Smith-Lovin and J. M. Cook (2001) Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology* 27: 415-444.

Miller, H. J. (2005) Necessary space-time conditions for human interaction. *Environment and Planning B: Planning and Design* 32: 381-401.

Mok, D., B. Wellman and R. Basu (2007) Did Distance Matter Before the Internet? Interpersonal Contact and Support in the 1970s. *Social Networks* Forthcoming:

Newman, M. E. J., D. J. Watts and S. H. Strogatz (2002) Random graph models of social networks. *Proceedings of the National Academy of Sciences of the USA* 99: 2566-2572.

Páez, A. and D. M. Scott (2004) Spatial Statistics for Urban Analysis: A Review of Techniques with Examples. *Geojournal* 61: 53-67.

Rieser, M., K. Nagel, U. Bueck, M. Balmer and J. Rümenapp (2007) Truly agent-oriented coupling of an activity-based demand generation with a multi-agent traffic simulation. Transportation Research Board Annual Meeting. Washington, D.C.

Rothenberg, R., S. Q. Muth, S. Malone, J. J. Potterat and D. E. Woodhouse (2005) Social and Geographic Distance in HIV Risk. *Socially Transmitted Diseases* 32: 506-512.

Shy, O. (2001) *The Economics of Network Industries*. Cambridge University Press, Cambridge

Silvis, J., D. Niemeier and R. D'Souza (2006) Social Networks and Travel Behavrio: Report from an Integrated Travel Diary. 11th International Conference on Travel Behaviour Reserach, Kyoto.

Watts, D. J. (1999) *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton Universtiy Press, Princeton