Towards a comprehensive agent-based simulation framework incorporating joint activity-scheduling and ride-sharing within households

Pieter J. Fourie, Future Cities Laboratory

Conference paper STRC 2012
Towards a comprehensive agent-based simulation framework incorporating joint activity-scheduling and ride-sharing within households

Pieter J. Fourie
Future Cities Laboratory
Singapore

Phone: +65 9116 8440
email: fourie@ivt.baug.ethz.ch
May, 2012

Abstract

Ride-sharing within households as the result of joint activity scheduling poses a challenging problem in transport demand modelling. This phenomenon is becoming increasingly important in environments such as Singapore, where extensive demand-management policies are implemented.

Various aspects of the problem have been addressed in isolation, but an integrated approach to full household activity schedule-based travel demand modelling including ride-sharing, with the aim of providing predictive capabilities for policy testing, remains elusive. In this paper, a framework is proposed for this capability to be integrated into the existing MATSim agent-based transport simulation system. Particular attention is paid to a possible hybrid simulation scheme, that could possibly allow a departure from complex, integrated meta-heuristic re-planning modules to simpler, atomistic modules that are easier to maintain. Such a simulation would then rely on long simulation runs for stable and realistic household activity patterns to emerge. The paper concludes with suggestions of possible atomistic re-planning modules that could be used in the incremental development of a full model of intra-household activity coordination.

Keywords

Intra-household coordination – ride-sharing – joint activities – agent-based simulation
1. Introduction

Intra-household decision making resulting in joint activities and ride sharing has received increasing attention since the turn of the century (Buliung and Kanaroglou 2007). However, due to the size and complexity of the problem, only parts of it have been examined in isolation. Operational models for the purpose of transport demand modelling therefore have to rely on a piece-wise approach to arrive at a user equilibrium of travel demand. One can therefore envision a system where one model is used for activity agenda generation (e.g. (Arentze and Timmermans 2004; Arentze and Timmermans 2009), another then needs to convert such agendas into activity schedules and locate them in space, another might have to ensure consistency in mode choice, and yet another performs network loading and adjustment of joint schedules to determine the impact on the transportation system (e.g. Dubernet, 2011).

Because systems aren’t integrated, each step in the demand modelling process relies on limited feedback based on the aspects of the problem that each individual model considers, and therefore diminishes our confidence in the realism of the household activity schedules that are produced, and their associated travel demand.

1.1 Application environment: Singapore

This paper proposes an integrated, multi-agent simulation-based approach to solving the problem of joint activities and household ride-sharing. The model will be developed with a view towards its application in Singapore, where it will form part of a larger research effort of applying state-of-the-art techniques to a functional agent-based model of Singaporean travel demand that includes the explicit modelling of road pricing, secondary activity location choice and public transport.

The Singaporean case highlights both the importance – and complexity – of modelling within-household coordination and joint travel. On the one hand, there are strict policy measures, such as highly restricted car ownership and peak period road pricing, and the availability of a complex mixture of modes. Furthermore, limited space and expensive real estate contribute to a relative prevalence of multi-generational households, with evolving roles and responsibilities for the members of said households.

Based on reported trips from the 2008 Household Interview Travel Survey (HITS), we see that, of the 9.8 million motorized trips made per day, 2.4 million are made by persons driving a private car, while 1.3 million trips are made by car passengers.
Table 1 shows that about one-third of these joint trips appear to have been made with someone from outside the household as driver. The modelling of ride-shares with drivers outside the household requires the modelling of social networks, which is considered to fall outside the scope of this study.

Table 1: Classification of joint trips in HITS

<table>
<thead>
<tr>
<th>Driver not in household (,000s)</th>
<th>Drop-off only (,000s)</th>
<th>Joint activity upon arrival (,000s)</th>
<th>Pick-up only (,000s)</th>
<th>Departure from joint activity (,000s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>421</td>
<td>284</td>
<td>113</td>
<td>239</td>
<td>244</td>
</tr>
</tbody>
</table>

The paper is structured as follows. Firstly, a survey of recent literature reveals the need for an integrated approach to the problem of joint activities and household ride-sharing. The agent-based simulation approach is suggested and conceptually illustrated in the following section. Particular attention is paid to a general scheme for accelerating the convergence of large-scale simulations that require the integration of many re-planning strategies in the face of complex supply options, as with the case of Singapore. A conclusion summarizes the suggested research agenda.
2. Literature survey

In their review article, Buliung and Kanaroglou (2007) note that the modelling of intra-household decision processes leading to activity-travel outcomes has only begun to receive attention at the turn of the century. A survey of the literature shows two streams of work; the first focusing on the understanding of household decision-making that leads to joint activities and ride-sharing, versus the application of said understanding to the generation of realistic household activity schedules.

In terms of understanding, several authors have followed random utility-based approaches, modelling household activity scheduling as a group utility maximization problem. However, because the choice set is of high dimension, different studies have focused on different dimensions of the problem.

Bradley and Vovsha (2005) classify activity travel diary data into a number of daily activity pattern (DAP) types. Their choice model considers all possible combinations of DAPs of all household members as alternatives, and estimates the contribution of both individual-level attributes and group-wise interaction terms representing the joint choice of the same pattern by several household members.

Gliebe and Koppelman (2005) also use the daily activity pattern as central concept in their model on combinations of two people in a household. They propose a structured discrete choice model that identifies a number of individual activity patterns that belong to ten joint outcome alternatives for the household. They find that significant predictors of pattern choices are commitment to work schedules, auto availability, commuting distance and the presence of children in the household.

Srinivasan and Athuru (2005) analyse within-household effects and between-household differences in the allocation of maintenance tasks to household members. They use a nested mixed logit model that decides first if an activity is performed jointly or individually, then selects the individual gaining highest utility for the activity if solo utility outweighs joint utility.

Srinivasan and Bhat (2006) develop a model that examines the utility of in-household/out of household joint/independent discretionary activities undertaken by household heads. Roorda et al. (2006) consider vehicle allocation, ride-sharing to joint activities, and pick-up & drop-off rides in a utility maximising framework to predict mode choice for entire households. Because of the vastness of the search space, they deploy a genetic algorithm approach on a
cluster of parallel operating computers to estimate the parameter set that maximises the log-likelihood function of a complex trip chain nesting choice structure.

The estimation and subsequent simulation of choice models, in order to produce household activity schedules for transport demand modelling, can only be applied to very limited cases, where both the size of the household and the alternatives open to it are very limited. Choice models require the enumeration of all possible alternatives, which is an unimaginably large number for any practical application. Therefore, when it comes to the generation of household activity schedules, literature generally reports the application of heuristic and meta-heuristic approaches.

Meister et al. (2005) extended the work of Charypar and Nagel (2005a) from the generation of individual to full household activity schedules. They use a genetic algorithm to generate individual schedules containing complete information of activity type, sequence, locations, start times, durations and connecting means of transportation. At the level of the household, their model considers division of work, joint activity participation and the allocation of means of transportation to household members. The fitness function used in their model evaluates household utility as the unweighted sum of the household individuals’ utilities. They modify the usual formulation of activity schedule utility beyond the individual by adding terms rewarding the degree of activity synchronization between individuals, and introduce a term reflecting the urgency of performing an activity based on the level of its need. They ran the model on a relatively small example, however, and note that several improvements need to be made for their model to be applied to large-scale scenarios.

Arentze and Timmermans (2004) use a step-wise approach of applying choice heuristics to produce complete household activity schedules. They use a CHAID decision tree induction method on activity diary data arrive at a decision tree that represents an exhaustive set of mutually exclusive rules for each decision step in the model. However, the model only produces a schedule for two adult persons in each household, and only considers activity allocation and joint activities performed by the two adults.

Charypar and Nagel (2005b) investigate the use of machine-learning approaches to activity schedule generation. Their model only considers the triple (type of activity, starting time of activity, time already spent at an activity) in determining whether to stay at an activity or transition to another activity. Their approach is developed further by Janssens et al. (2007) to include location information, non-restriction to a maximum number of activities and the incorporation of realistic travel times.

Arentze and Timmermans (2009) developed a model that employs a heuristic approach with feedback and learning to produce multi-day, multi-person activity agendas. Central to the
model is the concept of need, both at the person and household level, expressed in a utility-of-time threshold parameter that determines whether an activity is included in their schedule or not. The model does not, however, transform the activity agendas into travel demands, as information on activity location, exact timing, trip-chaining and transport mode is lacking. Märki et al. (2011) employ this concept of need in a multi-day agent-based simulation, where a need-based planning heuristic constructs activity schedules on-the-fly.

Dubernet (2011) applies the modelling of joint trips in MATSim, the multi-agent transportation simulation framework (Balmer et al. 2004; Balmer et al. 2009). This system simulates the day plans of each individual in the study area in a queue-based mobility simulation, and calculates the generalised costs of executed plans. Agents then add to their set of plans through a series of mutation algorithms applied to existing plans that are executed again in the mobility simulation. As the number of plans for each agent grows with increasing iterations, the worst performing plans are discarded, and the system gradually converges to a point corresponding to user equilibrium. In his modification, Dubernet takes a set of plans including joint trips for cliques of agents produced by an external system, and optimises the timing, joint trip participation, mode choice and routing of these plans using a genetic algorithm. Joint plans’ scores are taken as the sum of individual utilities that make up the plan, as interaction terms are not known for his generic cliques. Passengers in joint trips are not explicitly simulated; instead they are teleported according to expected travel time. This also means that they do not actually wait at activity locations for their shared ride to arrive. Dubernet employs a heuristic to adjust the departure times of clique participants such that the vehicle only departs when the last participant is present at the location.
3. Modelling intra-household coordination and joint trip making in MATSim

The MATSim framework is shown in Figure 1. The basic principle of operation is that a population of commuters with an initial demand of activity day schedules (plans) are generated, and executed in a mobility simulation that assigns their journeys to the transportation supply. The effectiveness of their performance is evaluated during a scoring phase, which assigns a utility to each performed plan, based on how much time was spent traveling and waiting versus actually being at the right place at the right time to perform an activity for a suitable duration. A re-planning phase takes existing plans and mutates them in order to come up with new plans. These are executed again and added to each agent’s set of existing plans. Once each agent reaches a maximum number of plans, the poorest performing plans are discarded. As a consequence, the population of plans gradually improves and approaches user equilibrium.

Figure 1 Principle of operation of the MATSim framework

The original design has travellers as individuals, with no interpersonal coordination. Dubernet (2011) improved on this concept with the introduction of cliques; groups of one or more persons that are able to travel together. This data structure will be applied to model intra-household coordination and keep track of members that can participate in joint trips and activities.

3.1 Best-response re-planning versus stochastic approaches

The original principle behind MATSim is that, through random mutation of plan elements, co-evolution of plans being executed in the same transportation network will converge to a solution that takes account of all pressures and interactions in the system. As long as there is (a) a mutation mechanism that allows an agent to respond to a particular pressure, like the ability to change activity timing and trip routing and (b) a scoring function that evaluates the
quality of the change, the system will eventually explore the solution space to such an extent that all agents should have come up with good solutions to their particular niche in the transport system ecology.

Initial implementations of MATSim relied on very simple atomistic mutation operations to operate on plans. For instance, in Balmer et al. (2005), two simple re-planning strategies are employed to change the timing and time-dependent routing of plans. Traffic count comparison shows how simulated volumes gradually adjust to congestion to give a close match to actual volumes.

Unfortunately, simple re-planning strategies require many iterations of the simulation in order to converge. These long computation times have served as motivation for the development of so-called ‘best-response modules’, such as those proposed by e.g. Meister et al. (2006). These approaches attempt to mutate more than one plan dimension simultaneously in a way that is anticipated to always result in an improvement in utility of the resulting plan.

These modules generally employ a meta-heuristic approaches, like genetic algorithms or tabu search. As an example, a simple genetic algorithm optimising activity timing would work as follows: starting from a base plan, a mutation and cross-over scheme generates a number of timing mutations for that plan, resulting in a population of plans for each agent. The expected performance of each resulting plan is calculated based on travel times from the previous iteration, using a utility function that rewards activity performance and punishes traveling and arriving late. Poorly performing plans are discarded and the process repeats until a highly improved version of the plan, given the constraints of the model, emerges. These modules require less simulation iterations to converge, but one might level a number of criticisms against them.

**Application specificity.** As a matter of necessity, meta-heuristics are purpose-built optimization techniques. For instance, if one employs a genetic algorithm (GA) to optimize for activity inclusion, order and duration, it requires the construction of a chromosome that presents those dimensions in a way that they can be acted upon by mutation and cross-over operators. If our requirements were to change to now include modelling individual mode choice, a complete overhaul of the module would be required, along with the construction of a new chromosome and mutation and cross-over operators.

Another alternative would be to have a separate mode choice allocation module to assign, say, random tour-based modes to connect activities in the schedules produced by the GA module. The resulting plans are then scored by using a utility function that includes mode-specific terms, along with the terms used in the GA. However, it should be apparent that such a two-step operation cannot guarantee a thorough search of the solution space. Specifically, the GA
approach will force the solution into areas favouring its dimensions in isolation, therefore only producing neighbouring solutions in the mode choice dimension.

In principle, one expects the combination of four atomistic operations (one adding or removing activities, another changing their order, another changing their duration and yet another changing our-based mode choice) to enter a domain of the search space that would balance the influence of all these operations in maximising plan utility.

**Super-optimisation.** Meta-heuristic modules tend to head straight for highly improved activity schedules (insofar as their dimensions of mutation allow), without passing through intermediate stages that reflect the stochasticity and lack of perfect knowledge in the actual transport system. As can be seen in Figure 2, departure times become sharp peaks aligned with facility opening and closing times, rather than the broader spread that we see for the case where random mutation was applied.

Figure 2 Comparison of departure times for random time mutation vs. best-response replanning

![Comparison of departure times](source: Meister et al. 2006)
Module maintenance and inter-operability. The MATSim platform is constantly evolving, with many collaborators working on various improvements and new capabilities. As new capabilities are added, certain re-planning modules need to be adapted to work with new data structures and program objects. For instance, the recent introduction of complete agent-based public transport simulation (Rieser 2010) necessitated a re-evaluation of the working of the so-called time allocation mutator module. This module would take an agent plan and randomly adjust the start time and duration of its activities within a 30 minute band. In order for it to work with the new transit simulation, it required the simple change of ignoring transfer activities when mutating activity durations. In comparison, integrated meta-heuristic re-planning approaches require a massive overhaul in order to have the same capability as it did before. The simple module required little insight to adapt to the change; the meta-heuristic module requires expert knowledge.

Furthermore, the complexity (and likeliness to break) of the model system as a whole is not necessarily a linear combination of the individual complexity of its re-planning modules. Atomistic re-planning modules can be turned on and off to determine the source of bugs. In comparison, the de-bugging of complicated, expert modules and their interaction with other modules can be a daunting task.

What tends to happen, therefore, is that expert modules are developed for the purpose of a specific project, and then rarely used again. One would clearly rather want a policy of re-usability and robustness to apply to the system design.

Simulation time. Increasing complexity necessarily means increased time required to perform a single iteration. The introduction of explicit public transportation modelling for our Singapore scenario increases simulation time from 10-15 minutes per iteration when transit agents are teleported to over 90 minutes for the full transit simulation. Similarly, more complex re-planning modules require more time to execute, but this sacrifice pays off in a smaller number of iterations for user equilibrium. In order to employ simpler re-planning modules, a general strategy for reducing simulation time needs to be formulated. In the next section, such a strategy is proposed. Then, a number of atomistic operations are suggested that could form the first modules in a process that will aim to model intra-household interaction, joint activities and trips with ever-increasing realism, in the presence of arbitrary model features such as public transport, road pricing and secondary activity location choice.

3.2 Hybrid simulation execution

The current implementation of the MATSim queue-based simulation allows one to simulate private vehicles and public transport explicitly, with other modes being teleported from start to end location using free-speed travel time multiplied by a mode-specific factor. The queue
simulation is multi-threaded, with simulation time decreasing with increasing threads. However, increasing the number of threads beyond a usual optimum of 3-5 increases simulation time again as threads need to communicate increasingly with each other in order to transfer vehicles between links allocated to different threads (Dobler 2010).

The queue simulation generates simulation events, such as when an agent enters or leaves a link, enters or leaves a public transport vehicle, starts or ends an activity, or waits to enter a link because that link is already at capacity, to name a few. Scoring functions ‘listen’ for these events in order to calculate plan performance on the fly. A simple scoring function, evaluating plan performance only on the basis of time spent at activities minus time spent traveling or arriving late, will only listen for a limited set of events, specifically activity start and end, and travel start and end events. These, along with the constraints on activity location opening and closing times, are enough to evaluate plan performance. A more complex scoring algorithm might additionally listen for link enter and leave events, in order to calculate dynamic road pricing components to its utility function.

If the number of threads optimal to a scenario really is a hard limit and the queue simulation cannot be further optimised, then the possibility of a hybrid system should be investigated. In principle, such a system will switch between the queue simulation and a highly simplified mobility simulation (let’s call it the Virtual Events Generating simulation – VEGsim), and let re-planning modules and scoring functions act on information from the last full mobility simulation. The principle of operation of such a hybrid system is illustrated in Figure 3.

Figure 3       Suggested operation of hybrid execution for MATSim
As an example, suppose a simple simulation where we only model private vehicle traffic, and only allow activity timing and routes to change through re-planning. We feed our simulation with a set of initial plans with na"ive timing and routing, and execute them in the simulation. Congestion causes a large number of plans to perform poorly. A number of plans are selected to have their activity timings randomly changed, and are re-routed through the network, using the dynamic link travel times from the previous iteration. Now, instead of using the queue simulation, we have the VEGsim module that simply generates link entry & exit and activity departure and arrival events based on link travel time information from the last full mobility simulation.

These events are processed in parallel using the usual events-listening scoring functions. The plan and its associated score is marked as ‘virtual’ and stored in memory with the rest of the agent’s plans. The simulation then proceeds to the re-planning phase, and once again, only acts on those plans that have been marked as virtual. The resulting virtual plans are once again executed in VEGsim, scored, and passed to the re-planning module. After a number of such iterations, a selection scheme is employed to select a virtual plan for execution in the full mobsim, while the rest of the virtual plans are discarded. This selection scheme can simply select the best virtual plan, or employ some stochastic mechanism based on virtual plan score. Following full execution, the process of plan selection, VEGsim execution, scoring and virtual re-planning repeats. The expectation is that, as it has been explained for this simple case, the operation of the planomat module proposed by Meister et al. (2006) has largely been replicated.

The way it is described here, the VEGsim module can be multi-threaded with no upper limit on number of threads, as there is no interaction with other vehicles. Any scoring function and associated re-planning modules can be used with this simplified simulation, as long as it generates the events they require. The module should also be relatively easy to extend to reproduce any new events generated by newer generation mobility simulation modules.

Data structures need to be put in place that allow for quick access to the previous full mobsim’s link travel times, and public transport travel times. As for the events that are generated, a quick-sorting data structure is required that can sort events from different threads into one stream, to be passed to an events processing module. Instead of generating events for all agents, those who have not been marked for VEGsim execution might have their events from the last simulation retained, and VEGsim events inserted into this events stream. Then the composite stream can be processed as a whole.
Essentially then, this mode of execution will allow the functionality previously afforded by best-response re-planning modules to be replicated and extended to apply to an arbitrary combination of re-planning modules and scoring functions.

### 3.3 Suggestions for atomistic re-planning modules to simulate intra-household coordination and joint trip-making

Essentially, the work of any of the authors that employ a utility-based approach, in the preceding literature survey section, can be converted into a scoring function for use in the agent-based simulation. Their methods for parameter estimation can be applied to the particular case being studied, in order to calibrate the scoring function to the case in point. Re-planning modules, on the other hand, can be made as simple as possible to facilitate fuller exploration of the solution space.

Implementation of these modules assume the existence of (at least) the data structures described by Dubernet (2011) to accommodate the concept of cliques, or in this case, households composed of agents. The agent description might also need to be expanded beyond its current implementation to include more descriptive demographic information.

For a start, one would want to test the conversion of Dubernet’s work into a series of simple atomic operators, before moving on to household maintenance activity allocation, and joint activity participation.

The first module one might propose would be a simple joint trip mutator. Suppose that a baseline plan for an agent that doesn’t have a vehicle or driver’s license only uses non-chain based modes like transit or taxi to travel between activity locations. A joint trip mutator module might evaluate the car availability and licensure of members of a household and simply change one of the qualifying agents’ mode of travel for an arbitrary trip to be a passenger of one of the licensed drivers. Such a module implies another module that is notified of joint trip generation, and ensures that, for instance, the driver agent’s plan is transformed to be at the passenger’s origin headed toward their destination at the right time.

A joint trip’s departure time should be allowed to evolve as with other independent activities. The time allocation mutator module therefore needs to be modified to maintain some form of synchronization between joint trips.

Another module might take a household’s plans, strip all maintenance and other discretionary activities currently allocated, then proceed to allocate such activities according to a modification of the scheme of Srinivasan and Athuru (2005), such that allocation of activities for the entire household are considered. The departure time of the joint activity could be
selected according to some decision rules that evaluate the timing of the fixed trips of the household. As the choice model approach cannot consider all location alternatives, one might allocate an arbitrary facility for a start; say the closest qualifying facility from the trip origin, or the qualifying facility closest to the centroid of the participants’ origins, if it’s a joint activity.

One would then chain this module with a secondary location choice model that would improve the secondary location choice based on prevailing link travel times and transit performance. The currently implemented secondary location choice model of Horni et al. (2008) would have to be adjusted to cope with joint activities – an initial implementation might look at the intersection of the time-use prisms of all activity participants.

With a view to quick simulation times, the passenger trip would not be explicitly simulated, but would always be teleported at the prevailing dynamic link travel times. The plan would be heuristically adjusted after simulation to have it not depart before the driver agent, and a scoring function similar to that of Meister et al. (2005) might be employed to reward the degree of synchronization in the executed plans.

At this stage, one wouldn’t want to suggest any further modules, because the limitations of such a conceptual system aren’t immediately apparent, and much development and testing would be necessary to come up with effective strategies. However, these simple examples hopefully illustrate how one might employ a gradual, incremental approach towards addressing the problem as a whole in an agent-based simulation framework.
4. Conclusion and outlook

An evaluation of the literature reveals the staggering complexity of the intra-household coordination, negotiation and decision-making processes that result in the observed activity patterns and trip-making of individual household members. A simulation-based approach towards the problem that would employ all currently implemented re-planning modules, while allowing for new modules that produce intra-household coordination and joint trips was proposed. This approach would rely on co-evolution and emergence to produce realistic schedules and an associated demand. As such, it requires quick simulation times in order to allow thorough exploration of a very large solution space. A possible approach for reducing simulation times and generalising best-response re-planning was proposed. A simple combination of re-planning modules was proposed to illustrate that, at the very least, the simulation-based approach would be able to produce and execute plans with joint trips.

It is suggested that such re-planning modules and their associated scoring functions can be developed in an incremental approach, where first one module is developed and its effect tested, followed by the development and testing of another, and then the joint deployment of multiple modules. Besides allowing one to investigate the influence of various module elements in isolation versus when they work together, an incremental approach allows one to investigate module interactions and isolate faults much easier.

First order of business would be to produce a proof-of-concept implementation of the proposed hybrid simulation system. If the system performs comparably to the best-response planning approach, it would warrant further investigation; both to reduce transit simulation times and to apply multiple re-planning strategies in combination.

If the hybrid simulation concept bears fruit, one would go ahead and develop re-planning modules that are as simple as they could be to produce the various elements that manifest during intra-household coordination that result in joint trip-making and joint activities.
5. Acknowledgement

This proposed study forms part of work funded by the National Research Foundation of Singapore.

The author thanks colleagues Marcel Rieser and Sergio Ordonez for their inputs and opinions.
6. References


