
New approaches to generating comprehensive all-day activity-travel schedules

Matthias Feil

Michael Balmer

Kay W. Axhausen

STRC 2009

August 2009

1 Abstract

Activity-based travel demand models derive travel demand from people's desire to pursue activities in time and space. They generate activity-travel schedules for individual travellers or homogeneous groups of travellers. Comprehensive activity-travel schedules hold information on which activities are performed, in which order, where and for how long, and which travel modes are used between the activities including corresponding routes. This paper presents PlanomatX, a new scheduling algorithm based on *Tabu Search* that generates comprehensively optimized all-day schedules. The paper furthermore presents a new concept of schedule recycling that significantly reduces simulation runtimes by re-using schedules of optimized travellers for other non-optimized travellers. Both PlanomatX and schedule recycling are part of the agent-based microsimulation MATSim (Multi-Agent Transport Simulation, <http://matsim.org>). MATSim's utility function has been adapted to cope with the enhanced functionality of PlanomatX and schedule recycling. First test results on the greater Zurich scenario with more than 170,000 agents show that PlanomatX achieves significantly better optimization results than MATSim's existing scheduling algorithms. However, it also leads to disproportional simulation runtimes. Schedule recycling relieves this drawback and allows for generating comprehensively optimized all-day schedules for large-scale scenarios at affordable runtimes.

2 Introduction

Activity-based travel demand models emphasize travellers' participation in out-of-home activities as a source for travel demand. They understand travel demand as a derived demand that arises from people's desire to pursue activities in time and space (Oi and Shuldiner, 1962). In comparison with classic transport models, activity-based models allow for an increased recognition of the complexity of travel decisions, based on a more behaviourally sound model framework (Cirillo and Axhausen, Forthcoming). Particular advantages are an improved capability to model non-work and non-peak travel, an improved capability to move beyond traditional explanatory variables (i.e. zone-based socio-economics, travel time and cost), and an improved capability to deal with the effects of household interaction, age, lifestyle, etc. on travel behaviour (Meyer and Miller, 2001).

Activity-based travel demand models generate activity-travel schedules for individual travellers or homogeneous groups of travellers. Comprehensive activity-travel schedules hold information on which activities are performed, in which order, where and for how long, and which travel modes are used between the activities including corresponding routes. The fundamental

problem of scheduling is the combinatorial size of feasible outcomes. The scheduling process of a traveller may easily exceed millions of alternatives (Bowman and Ben-Akiva, 1996) so that the problem is generally not solvable by complete enumeration. Travellers do not perceive the immense magnitude of the solution space since they simplify their decision making process. They consider just a few discrete alternatives they choose from. Modellers' task is to develop concepts that are capable of handling the immense solution space with acceptable computational efforts but also in a way that matches the behaviour of the travellers.

We propose a new scheduling algorithm *PlanomatX* that generates comprehensively optimized all-day schedules, i.e. optimal combinations of a schedule's activity chain (number, type and sequence of activities), activity timings, and the location, mode and route choices. Furthermore, we present the new concept of *Schedule Recycling* that significantly reduces simulation runtimes by re-using schedules of optimized travellers for other non-optimized travellers. Both *PlanomatX* and schedule recycling are part of the agent-based microsimulation MATSim (Multi-Agent Transport Simulation Toolkit, MATSim-T, 2008). MATSim's utility function has been adapted to cope with the enhanced functionality of *PlanomatX* and schedule recycling.

This paper first presents related work on the problem of activity-travel scheduling followed by an introduction to MATSim and its core principles. Then, the new scheduling algorithm *PlanomatX* is described including necessary adjustments of MATSim's utility function and a presentation of empirical results. Furthermore, we also illustrate the concept of schedule recycling. The paper closes with a summary and outlook.

3 Related work

Scheduling in activity-based travel demand modelling follows three major lines of research (Ettema and Timmermans, 1997; Goulias, 2002; Timmermans, 2001, 2003):

- **Econometric models** "... use systems of equations to capture relationships among attributes" (Bhat *et al.*, 2004). They are disaggregate and yield probabilities of choices. The probabilities may be translated into specific schedule solutions through Monte-Carlo simulation. Recent contributions in the field of activity-based travel demand modelling are, for instance, Bhat (1997, 1998), Bowman and Ben-Akiva (2001), and Habib and Miller (2008). Moreover, the journal of *Transportation* dedicated a Special Issue to modelling intra-household interactions and group decision making (Bhat and Pendyala, 2005) featuring several econometric models (Srinivasan and Athuru, 2005; Srinivasan and Bhat, 2005; Bradley and Vovsha, 2005; Gliebe and Koppelman, 2005). Econometric models

bear the advantage that they are based upon a well-established statistical methodology and economic theory. However, they tend to quickly become very complex and difficult to estimate and operationalize. Modellers are required to apply levers like limiting the level of detail of the model outcome (e.g., low temporal granularity, implicit location choice, etc.), or neglecting temporal, institutional, or spatial constraints (Joh, 2004).

- **Utility-based microsimulations** apply a sequential decision making process. Employing complex search heuristics they iteratively narrow down the solution space. Rather than a probability distribution, the result is always a precise solution alternative. Significant contributions in the field of activity-based travel demand modelling are e.g., STARCHILD (Recker *et al.*, 1986a,b), ORIENT (Axhausen, 1988), PCATS (Kitamura, 1996; Kitamura and Fujii, 1998; Pendyala *et al.*, 2005), CEMDAP (Bhat *et al.*, 2004), TRANSIMS (Hobeika, 2005; TRANSIMS, 2009), and MATSim (Balmer, 2007; Meister *et al.*, 2009; Balmer *et al.*, 2008a). Also the works of Charypar and Nagel (2005) and Meister *et al.* (2005) fall into this line of research. Based upon the iterative procedure and the intensified employment of heuristics, utility-based microsimulations are capable of reducing the complexity of econometric models. Moreover, they explicitly model the variability of travel demand (“emergent behaviour”) rather than probability-based average values, allow to explicitly include constraints along the sequential decision making process, and allow to take into account “competition” that arises from modelling the constraints (Vovsha *et al.*, 2002).
- **Computational process models (CPMs)** try to overcome the drawback of utility-based models, namely that travellers do not make “optimal” decisions but rather context-dependent heuristic decisions (Joh, 2004). CPMs “... *replace the utility maximising framework with behavioural principles of information acquisition, information representation, information processing, and decision making*” (Golledge *et al.*, 1994). CPMs are basically also microsimulations due to their disaggregate nature, the sequential decision process and the use of heuristics. However, the heuristics employed by CPMs rather consist of “if-then” rules than utility-maximizing decision criteria. Recent models in this line of research are SCHEDULER (Golledge *et al.*, 1994), AMOS (Pendyala *et al.*, 1995; Kitamura, 1996; Kitamura and Fujii, 1998), and ALBATROSS (Arentze and Timmermans, 2004).

Both the scheduling algorithm PlanomatX and the concept of schedule recycling break new ground. PlanomatX does so because it generates comprehensively optimized all-day schedules for large-scale scenarios. The above models either tackle only partial scheduling problems (e.g., choice of post home-arrival activity participation (Bhat, 1998), maintenance activity allocation among household members (Srinivasan and Athuru, 2005), and so forth), or are comprehensive all-day schedulers but produce very low temporal resolution (Bowman and Ben-Akiva, 2001) or

work on fictive small data sets (Charypar and Nagel, 2005; Meister *et al.*, 2005). The schedule recycling concept breaks new ground because it presents a way to re-use schedules of travellers for other travellers. No such concept is known to the authors in the literature.

4 MATSim overview

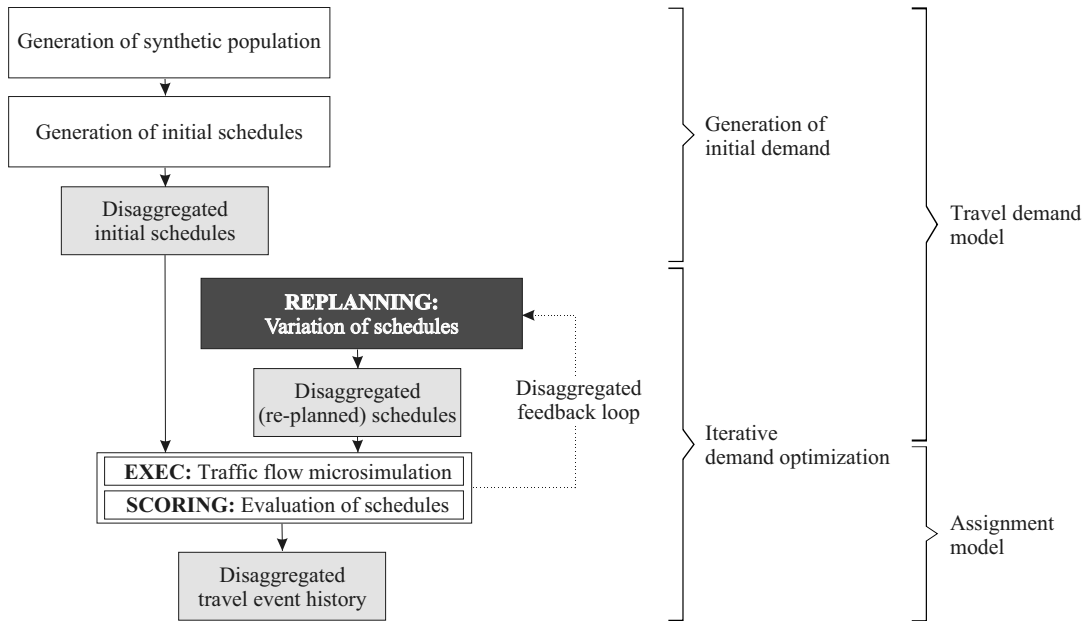
PlanomatX and schedule recycling extend the utility-based microsimulation MATSim (Multi-Agent Transport Simulation, Balmer, 2007; Meister *et al.*, 2009; Balmer *et al.*, 2008a). MATSim implements an activity-based approach to travel demand generation for large samples. Unlike other transport simulation models MATSim is agent-based throughout and produces individual activity schedules as input to the traffic flow simulation rather than origin-destination matrices as typically used in dynamic traffic assignment (Illenberger *et al.*, 2007). Initial demand schedules are generated by disaggregating census data. The schedules are executed and overall travel costs calculated using a suitable traffic flow microsimulation. Henceforward, the utility of the schedules is iteratively improved against the background of overall travel costs (see figure 1). Central to the improvement of the schedules is MATSim's replanning step where agents are allowed to learn and optimize their schedules. Since always a certain share (e.g., 10%) of the overall set of agents do so simultaneously MATSim's simulation process is a co-evolutionary learning process. The co-evolutionary learning process stops when none of the agents can further improve their schedule, or at an externally given maximum number of iterations. MATSim is developed jointly by TU Berlin, ETH Zurich and CNRS Lyon. It has been applied to several scenarios such as Switzerland, Berlin-Brandenburg/Germany, Toronto/Canada, Padang/Indonesia, among others.

MATSim's existing replanning step features algorithms to optimize the location choices (Horni *et al.*, 2008), the route choices (Lefebvre and Balmer, 2007), and the mode choices together with the activity timings (Balmer *et al.*, 2008b). Both PlanomatX and schedule recycling integrate into MATSim's replanning step. PlanomatX enhances the replanning step with capability to add the structure of the activity chains as a further dimension to the co-evolutionary learning process. Schedule recycling enhances the replanning step reducing the need to optimize each agent individually.

5 PlanomatX: Comprehensive schedule optimization

PlanomatX generates comprehensively optimized schedules, i.e. optimal combinations of a schedule's activity chain (number, type and sequence of activities), activity timings, and the

Figure 1: MATSim structure: PlanomatX and schedule recycling being new algorithms of the replanning step.



location, mode and route choices.

5.1 Problem formulation

The optimality, or fitness, of a schedule is measured against its utility. Optimizing a schedule means to maximize its utility. In MATSim, the problem of maximizing a schedule's utility is expressed by the following objective function, subject to the activity types' opening times (see Figure 3a):

$$\max U_{total} = \max \left[\sum_{i=1}^n U_{perf,i} + \sum_{i=1}^n U_{late,i} + \sum_{i=1}^n U_{travel,i} \right] \quad (1)$$

where U_{total} is the total utility of the given schedule; n is the number of activities/trips; $U_{perf,i}$ is the (positive) utility gained from performing activity i ; $U_{late,i}$ is the (negative) utility gained from arriving late at activity i ; and $U_{travel,i}$ is the (negative) utility gained from travelling trip i . $U_{perf,i}$ is a log function. $U_{late,i}$ and $U_{travel,i}$ are linear functions.

The objective function formulates a mixed-integer, non-linear, non-convex problem. An analytical, hence quick, solution approach is unknown for this class of problems. It may thus "... not be possible to solve for an optimal solution. [...] It still is important to find a

good feasible solution that is at least reasonably close to being optimal. Heuristic methods commonly are used to search for such a solution” (Hillier and Lieberman, 2005).

Moreover, the current utility function for the performance of activities is problematic. Its log form leads to unrealistic results when we allow for changes in the number of activities in the schedule. Due to the decreasing marginal utility, the log form would lead to schedules with a lot of very short activities. In other words, a schedule of two 30 minutes activities of a certain type would always be better than a schedule of once 60 minutes of the same activity. We therefore need a utility function for the performance of activities that can cope with a flexible number of activities in the schedule.

5.2 Heuristic solution algorithm

Charypar and Nagel (2005) and Meister *et al.* (2005) have shown that a Genetic Algorithm (GA) is able to solve the above stated utility maximization problem. Their work has led to the implementation of a GA that optimizes the mode choices and the activity timings of a schedule (see Balmer *et al.*, 2009, for more details). The extension of this algorithm towards a comprehensive schedule optimization would be a straight-forward development path. Yet “GAs are known as rather inefficient” (Charypar and Nagel, 2005). They relax very robustly to a (nearly-)global optimum (e.g., Hillier and Lieberman, 2005; Hasan *et al.*, 2000) but risk spending a significant part of their search process in potentially unpromising areas of the solution space. One reason is the missing cycling prevention, another the random mutation operator (Rahoual and Saad, 2006). Considering MATSim’s requirements, a perfect schedule optimization is desirable but not imperative. In fact, in many cases what people use as their schedules “... is far from being optimal” (Charypar and Nagel, 2005). The schedule optimization rather needs to be “good” but not optimal, as long as the computational time is kept low.

The class of Gradient Algorithms matches the above requirements. From any initial solution, the Hill Climbing Algorithm just follows the steepest gradient of the objective function. It stops when no further improvement is possible. It relaxes quickly but the solution is likely to be a local optimum. The Tabu Search Algorithm (Glover, 1989) is a more elaborate Gradient Algorithm that overcomes this drawback. It is equal to a Hill Climbing Algorithm until it has found the first (local) optimum. It may then select inferior solutions till a better solution has been found. A tabu list avoids cycling. The list stores the selected solutions of the t last iterations. Moves are forbidden that would reach these solutions again. A stop criterion (e.g., overall number of iterations or minimum improvement over n last iterations) lets the

algorithm finish. The Tabu Search algorithm quickly reaches “ok”-solutions followed by gradual improvement steps.

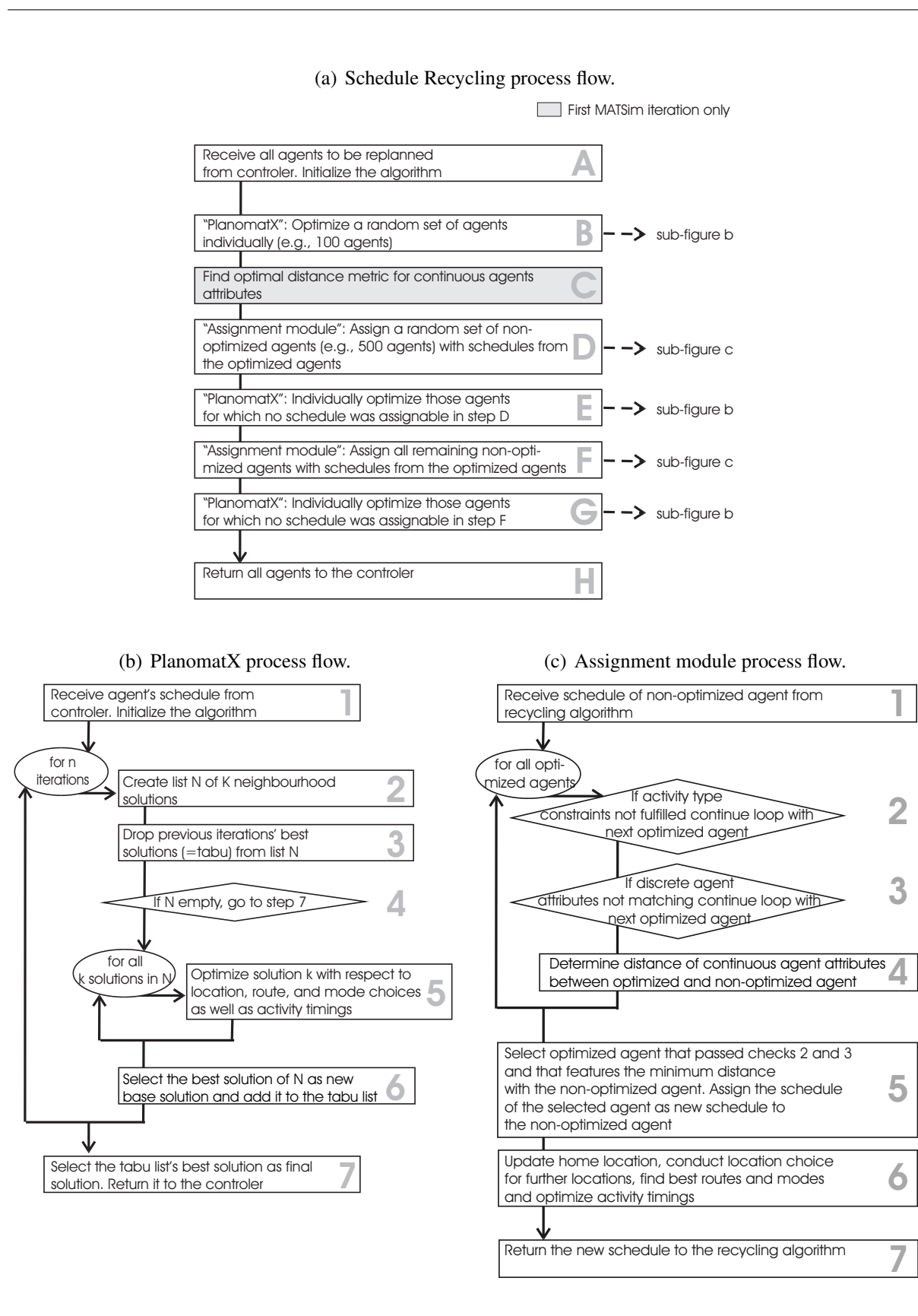
PlanomatX implements a Tabu Search heuristic. Given an arbitrary base solution, PlanomatX tries to proceed towards the steepest gradient, or in more descriptive words, towards the steepest ascent of the utility mountain range. Considering MATSim’s scheduling problem, the utility mountain range has not only two dimensions (north-south, east-west) but five: the activity chain, the location, route and mode choices, and the activity timings. In order to master the increased number of dimensions the PlanomatX algorithm solves the problem hierarchically applying two nested optimization loops (see figure 2b):

- *Outer loop, steps B/C/E*: The outer loop implements the Tabu Search principle and solves for the best activity chain. The loop steers the creation of a list N of K neighbourhood activity chain solutions (number, type and order of activities), drops those neighbourhood solutions that are tabu, scores the remaining neighbourhood solutions, selects the best from among them, updates the tabu list with the best solution, and sets the best solution as next iteration’s base solution. The activity chain neighbourhood solutions are created from this base solution through insertion/deletion of an activity, change of the sequence of activities, or change of the type of an activity.
- *Inner loop, step D*: For each created neighbourhood activity chain k , the inner loop optimizes the lower tier decisions of location, route and mode choices as well as of the activity timings.

The nested loop structure facilitates the integration of MATSim’s existing replanning algorithms dealing with the optimization of location, route, mode choices and activity timings. Yet more importantly, the nested loop structure implies that, when choosing the best activity chain, the outer loop in fact chooses the best schedule that is possible for the activity chain. Hence, no neighbourhood solution requires in future outer loop iterations to be investigated for different combinations of location, route, mode choices or activity timings. This quickly limits the solution space and is a considerable reduction of complexity.

The algorithm stops when an externally given maximum number n of outer loop iterations has been reached or when no more non-tabu neighbourhood solutions are available. The algorithm’s optimization result is the highest-scored solution of the final tabu list.

Figure 2: Process flows of the newly developed algorithms.



5.3 Utility function

PlanomatX measures the optimality of a schedule against its utility. MATSim’s existing utility function features a log form that, in combination with PlanomatX, would lead to schedules with lots of very short activities due to the decreasing marginal utility of the log-form. What we require is a utility function for the performance of activities that formulates some sort of optimal activity duration by its functional form. Assuming an average value of time, the utility function should feature segments where its value of time is below the average value of time, and segments where it is above. The optimal activity duration will be found in the latter segments. Joh (2004) presents a utility function matching these requirements:

$$U_{perf,i}(t_{perf,i}) = U_i^{min} + \frac{U_i^{max} - U_i^{min}}{(1 + \gamma_i \cdot \exp[\beta_i(\alpha_i - t_{perf,i})])^{1/\gamma_i}} \quad (2)$$

The function is an asymmetric S-shaped curve with an inflection point, originally developed in biological science (see figure 3; parameters set trying to match MATSim’s existing average values of time).

5.4 Application tests and results

We applied the PlanomatX algorithm to the Greater Zurich Scenario (Balmer *et al.*, 2009). The scenario comprises a set of 172,598 agents. They are a 10% random draw from those agents whose initial routes cross a 30 km circle around Zurich’s city centre. The road network is represented by a model network of 60,000 directed links and 24,000 nodes. There are 1.3 million home locations and more than 380,000 out-of-home locations. We assume car driver, public transport, and walk as the available transport modes. Ten activity types are modelled (two work types, five education types, shopping, leisure, and home). Home must always be the first and last activity type. In the initial demand generation, agents may have been associated with further agent-specific “primary” activities (e.g., work at facility x). The primary activities must be performed during the day. The agent is not allowed to drop them from its schedule but the duration and the schedule position of the primary activities are still flexible.

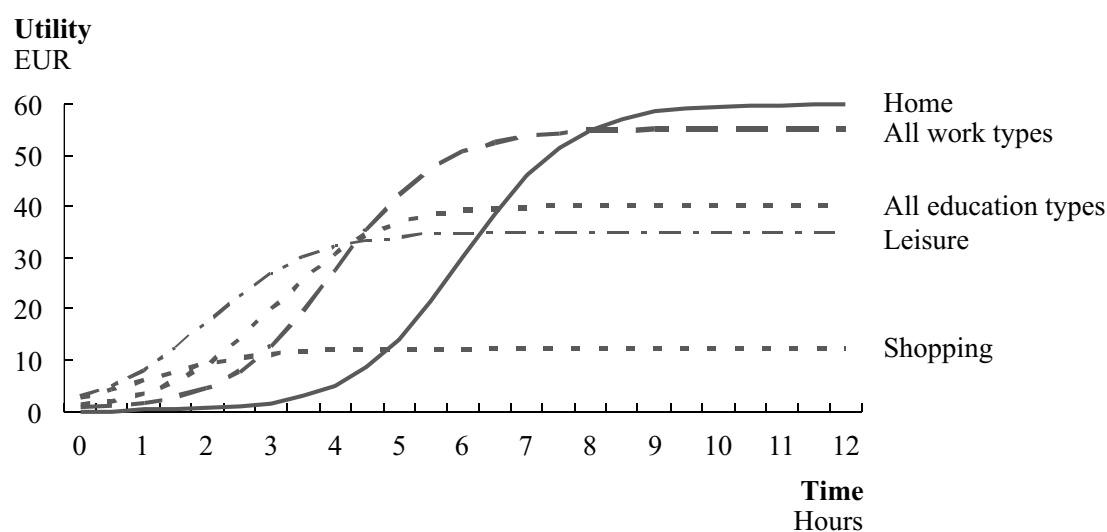
For all tests in this and the following chapter, we ran 50 MATSim iterations. Following an extensive parameter analysis (Feil, Forthcoming), PlanomatX was set to 20 outer loop iterations and a neighbourhood size of 10 solutions. Creating the neighbourhood solutions, shares of 30% insertion/30% deletion of an activity, 20% change of the sequence of the activity chain, and 20% change of the type of an activity were applied.

Figure 3: New utility function for the performance of activities (adopted from Joh, 2004).

(a) Parameters and constraints of the utility function.

Activity type	Parameters					Constraint
	U_i^{min} (EUR)	U_i^{max} (EUR)	α_i	β_i	γ_i	Opening times
Home	0	60	6	1.2	1	00.00 - 24.00
All work types	0	55	4	1.2	1	08.00 - 18.00
All education types	0	40	3	1.2	1	18.00 - 22.00
Leisure	0	35	2	1.2	1	07.00 - 16.00
Shopping	0	12	1	1.2	1	10.00 - 18.00

(b) Plot of the utility function.

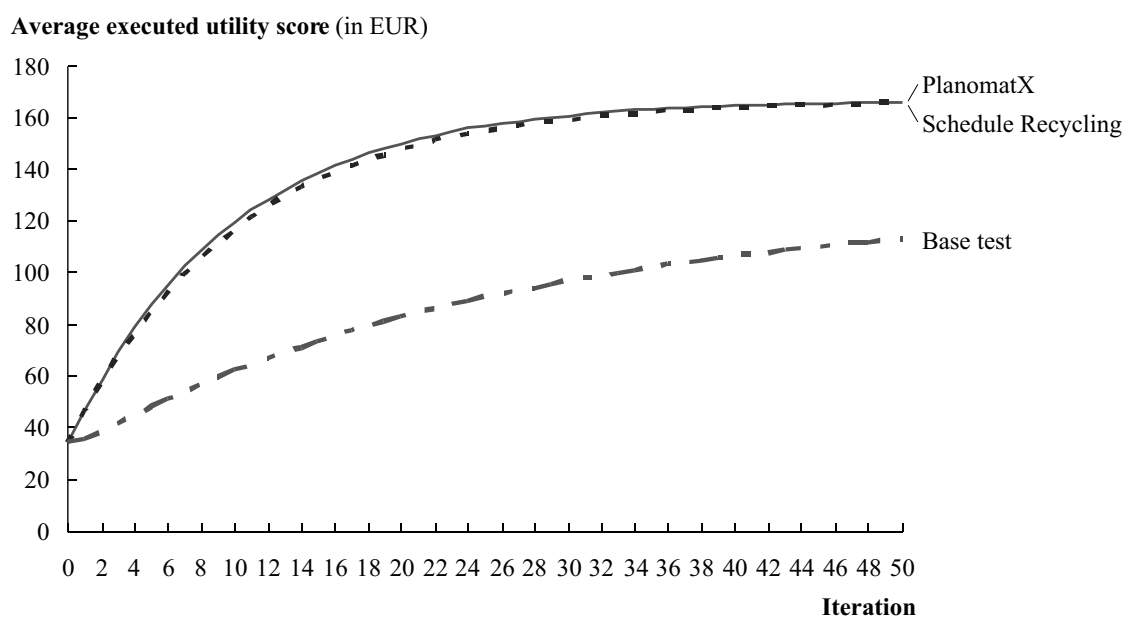


The following paragraphs demonstrate PlanomatX's optimization impact. It is displayed against a base test with MATSim's existing replanning algorithms, i.e. optimization of activity timings and mode, location, and route choices (c.p.). Hence, the base test involved all replanning dimensions of PlanomatX but the activity chain dimension.

5.4.1 Utility scores

Figure 4 displays the development of the average utility score of all agents' executed schedules. After 50 iterations, the base test attains a score level of 113 EUR. PlanomatX reaches a level of 166 EUR. This is a plus by 47% against the base case. It becomes evident that the additional activity chain dimension has a major impact. This is reasonable when we look at agents' optimized schedules.

Figure 4: Development of average executed utility scores for PlanomatX, schedule recycling and base test.



5.4.2 Optimized schedules

The improvement of the utility scores can be tracked in the schedules of the agents. Columns 1 and 2 of table 1 provide a macroscopic view of agents' activity chains as a result of the base test and of the PlanomatX test. In the base test, the activity chains are the initial demand chains drawn from the census data. The activity chains are identical before and after the optimization since none of the algorithms used in the test modifies the activity chains. Altogether, 2,762 different activity chains exist. The 20 most frequent activity chains account for 97,172 (56.3%) of the 172,598 agents in the scenario. The average length of the activity chains is 4.92 activities. For the PlanomatX test, the activity chains change from the initial demand chains (see base test) to the optimized chains as displayed in table 1. Now, the overall number of different activity chains is 652 indicating that PlanomatX reduces the overall diversification of activity chains. The 20 most frequent activity chains account for 156,124 (90.4%) of the 172,598 agents in the scenario. The average length of the activity chains has increased to 5.38 activities.

Drilling down to a microscopic view, figure 5 shows, as an illustration, the initial demand schedule and the optimized schedules of sample agent 1000914. In the base test, the score has improved from 20.45 EUR to 41.10 EUR. The initial demand activity chain "home-work_sector3-home-work_sector3-home" has, by definition, remained stable. The activity timings have slightly changed (e.g., the agent returns home at 6:49pm rather than at 7:40pm). The agent continues using public transport. In the PlanomatX test, the utility score has risen to 169.47 EUR. This is a plus of 312% compared to the base test. The activity chain has

Table 1: Activity chain statistics for PlanomatX, schedule recycling and base test.

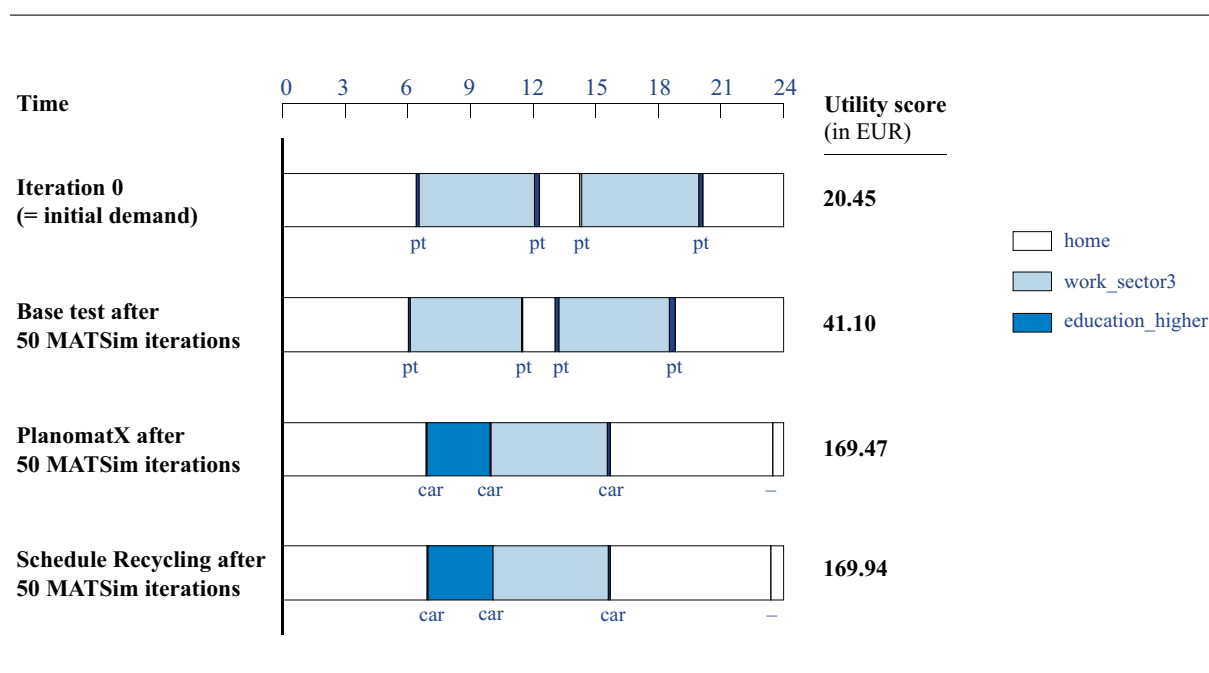
Base test (= initial demand)		PlanomatX		Schedule Recycling		Activity chain								
Ranking	Number of agents	Ranking	Number of agents	Ranking	Number of agents									
1	15500	61	103	46	91	home	work_sector3	home						
2	14122	73	70	48	79	home	leisure	home						
3	11265	80	59	55	62	home	shop	home						
4	8954	94	43	63	42	home	work_sector2	home						
5	7130	95	42	64	39	home	education_primary	home	education_primary	home				
6	5317	107	34	75	28	home	work_sector3	home	work_sector3	home				
7	3660	153	15	76	25	home	work_sector3	home	leisure	home				
8	3340	132	21	95	16	home	education_primary	home						
9	3218	137	20	87	17	home	work_sector3	leisure	work_sector3	home				
10	3213	153	15	75	28	home	work_sector2	home	work_sector2	home				
11	3126	134	20	87	17	home	shop	home	leisure	home				
12	2968	156	14	106	13	home	leisure	home	leisure	home				
13	2423	140	18	97	15	home	work_sector3	shop	home					
14	2256	164	12	116	10	home	education_primary	home	education_primary	home	leisure	home		
15	2094	6	10813	6	12690	home	education_primary	home	leisure	home				
16	2066	182	8	108	12	home	work_sector2	home	leisure	home				
17	1914	169	11	75	28	home	work_sector2	leisure	work_sector2	home				
18	1769	158	13	108	12	home	work_sector3	leisure	home					
19	1421	221	5	124	9	home	leisure	leisure	home					
20	1416	211	6	99	14	home	work_sector3	home	shop	home				
-	-	1	24790	1	25004	home	work_sector3	shop	shop	home	home			
-	-	2	22624	2	23570	home	work_sector3	shop	home	home				
-	-	3	21737	3	23207	home	work_sector2	shop	shop	home	home			
-	-	4	14899	4	14918	home	education_other	work_sector3	home	home				
-	-	5	12049	7	12377	home	work_sector2	shop	home	home				
-	-	7	10057	5	13367	home	education_other	work_sector2	home	home				
-	-	8	5845	10	4776	home	education_higher	work_sector3	home	home				
-	-	9	5249	8	6037	home	home	home	home					
-	-	10	4441	13	2748	home	education_primary	shop	shop	shop	home	home		
-	-	11	4201	11	3496	home	home	work_sector3	shop	home				
-	-	12	3429	15	1682	home	home	work_sector2	shop	shop	home			
-	-	13	3251	9	5976	home	education_higher	work_sector2	home	home				
-	-	14	3127	16	1380	home	home	work_sector3	shop	shop	home			
-	-	15	2611	26	738	home	home	work_sector2	shop	home				
-	-	16	2223	12	3349	home	education_secondary	work_sector3	home	home				
-	-	17	1735	14	2327	home	education_secondary	work_sector2	home	home				
-	-	18	1134	17	1334	home	work_sector3	home	home					
-	-	19	1002	22	763	home	education_primary	shop	shop	home	home			
82	245	20	907	27	662	home	education_kindergarten	home	leisure	home				
	97172		156124		160785	Subtotal of agents for 20 most frequent activity chains								
21 - 2762	75426	21 - 652	16474	21 - 457	11813									
	172598		172598		172598	Total of agents in scenario								
	4.92		5.38		5.30	Average length of activity chain								

changed to “home-education_higher-work_sector3-home-home” (the last short home activity can be considered as the start of the following day’s home activity). In light of the different sequence of activities, all activity timings have changed, and the agent now uses the car (no mode for last leg since the two home activities take place at the same facility).

5.4.3 Runtime

Both PlanomatX and base test were run on a high-end computer with 16 Itanium-2 Dual Core processors. The overall runtime of the PlanomatX run was 11 days and 11:23 hours. Each MATSim iteration took about 5:30 hours splitting into 15 minutes of traffic flow simulation and 5:15 hours of PlanomatX replanning (1.095 seconds per agent). The overall runtime of the base test was only 18:48 hours. Here, each MATSim iteration took about 22:30 minutes splitting into 15 minutes of traffic flow simulation and 7:30 minutes of replanning runtime

Figure 5: Initial demand schedule and optimized schedules after base test, PlanomatX and schedule recycling for sample agent 1000914 (for the ease of illustration, locations and routes are not shown).



(0.026 seconds per agent). One must conclude that PlanomatX's better scoring results come at a high runtime cost: While the scoring results have improved by 47%

- the overall runtime has increased by 1,350%, and
- the pure replanning runtime has increased by 4,100%.

5.5 Discussion

PlanomatX is a new Tabu Search algorithm that comprehensively optimizes activity-travel schedules. It extends MATSim's replanning functionality. Tests on the Greater Zurich Scenario have demonstrated PlanomatX's optimization performance. In comparison with the joint application of MATSim's existing replanning strategies (base test), PlanomatX reaches 47% higher utility scores. This is possible since the structure of the activity chains is now a dimension of the co-evolutionary learning process, too. However, in its current form, PlanomatX requires disproportional runtimes (+1,350%/+4,100%). Aiming for reduced runtimes, the following chapter will introduce a concept of schedule recycling that avoids running PlanomatX for each agent individually.

6 Schedule recycling

Schedule recycling is a concept to re-use schedules of optimized agents for non-optimized agents. Aiming for reduced runtimes, it avoids running PlanomatX for each agent individually.

6.1 Problem formulation

PlanomatX successfully optimizes activity-travel schedules but leads to disproportional runtimes. A major reason for the heavily increased runtimes is the activity chain being a new dimension of MATSim's co-evolutionary learning process. When analyzing table 1 evidence can be found that many agents share the same or similar activity chains even if the activity chain is flexible. A set of 652 different activity chains is sufficient to cover the activity chains of the scenario's more than 170,000 agents. The 20 most frequent activity chains already cover more than 150,000 agents. Aiming for reduced runtimes, this motivates to establish a way to re-use, or "recycle", schedules of optimized agents for other non-optimized agents, without running PlanomatX for each agent individually. The following problems arise: How may a process flow of schedule recycling look like? Which elements of a schedule may be recycled? According to which rule, or metric, agents may be assigned with those elements?

6.2 Solution algorithm

The schedule recycling algorithm (see figure 2a) is an alternating sequence of optimizing agents individually ("PlanomatX") and assigning the activity chains of optimized agents' schedules to non-optimized agents ("Assignment"). The core innovation of the algorithm is the assignment module (steps D and F). For a non-optimized agent, the module selects the best matching schedule from a set of individually optimized agents, assigns the non-optimized agent with the activity chain of the schedule and finalizes the assigned schedule with regard to location, route, and mode choices as well as activity timings. Figure 2c focuses on the process flow of the assignment module. The selection of the best matching schedule comes in three stages:

- *Check of activity type constraints (figure 2c, step 2)*: This first check verifies whether the schedule of the optimized agent contains all primary activity types of the non-optimized agent. This check also verifies whether the schedule of the optimized agent does not contain activity types that are not eligible to the non-optimized agent. An example would be "work" for a 10-years old child or "primary education" for a 50-years old adult.
- *Check of discrete agents attributes (figure 2c, step 3)*: Some agent attributes may prevent an agent from receiving a recycled schedule from another agent. For instance, an agent

without driver's license may not adopt a schedule containing car trips. One may also define that a schedule of a male agent may only be recycled for another male agent, and so forth.

- *Distance of continuous agents attributes (figure 2c, step 4)*: From the set of individually optimized agents, it is likely that several schedules will pass the first two checks. This third check determines which of these schedules matches the non-optimized agent best. It refers to the distance, or similarity, between the continuous attributes of the optimized agent i and the non-optimized agent j . The distance d is calculated as follows:

$$d_{i,j} = \sum_{k=1}^n |x_{i,k} - x_{j,k}| \cdot \delta_k \quad (3)$$

where δ_k is the weight coefficient of the continuous attribute k . Common continuous attributes may, for instance, be the agents' geographic distances between their primary activities (e.g., home-work-home) or their age. The definition of the multidimensional distance metric $\Delta = \{\delta_1; \dots; \delta_k; \dots; \delta_n\}$ will be described below.

Having iterated through all optimized agents, the assignment module selects the schedule that passed the two first checks and whose agent features the highest similarity with the non-optimized agent. It assigns the non-optimized agent with that schedule, i.e. it copies the activity chain of the schedule and pastes it into the new schedule of the non-optimized agent. It then updates the home location, conducts a location choice for the further activities, finds best routes between the locations, and optimizes activity timings and mode choices. We can see that, for an agent that is assigned with a recycled schedule, location, route, and mode choices and the optimization of the activity timings need to be conducted only once. This is opposite to PlanomatX where they need to be conducted for each of the agent's neighbourhood solution of the Tabu Search.

Let us highlight the role of the three stages of selecting the best matching schedule: The first two checks are responsible for the alternating sequence of optimizing agents individually ("PlanomatX") and assigning schedules of optimized agents to non-optimized agents ("Assignment"). This is because an a-priori search for optimized schedules passing the two checks for all non-optimized agents would be very cumbersome. It is much easier to just select some agents randomly, optimize their schedules individually (step B) and see whether these schedules pass the checks for the non-optimized agents (step F). Some non-optimized agents will remain for which no schedule is assignable. They need to be optimized individually (step G). This sequence would perfectly work but our algorithm refines it introducing steps D and E. Step D tests for a limited number of non-optimized agents whether the optimized schedules of step B pass the checks for these non-optimized agents. Those non-optimized

agents for which no schedule is assignable are handed over to step E. Step E optimizes them individually and adds them to the list of optimized agents of step B. Like this, they are available for step F. Step's F assignment now produces considerably less non-optimized agents for which no schedule is assignable. As a consequence, the runtime of step G is shortened, and the runtime reduction is very likely to be much higher than the runtime increase from step E. In theory, one may introduce more instances of steps D and E. In the most advanced case, every non-optimized agent for which no schedule is assignable is immediately optimized individually and added to the list of optimized agents. However, this makes parallelization of the algorithm extremely hard. We feel that the current setup of the algorithm is a good trade-off (Feil, Forthcoming).

Distance metric Δ of check 3 is central to the quality of the assignment module. In fact, check 3 conducts some sort of cluster analysis. It clusters the non-optimized agents around the (assignable) optimized agents and, in each cluster, the non-optimized agents are assigned with the schedule of the optimized agent. In conventional cluster analysis, the distance metric Δ is given (Euclidean distance, Hamming distance, etc.) and attention is paid to the most efficient way of clustering. In our case, the way of clustering is quite straight-forward (see above) but the distance metric Δ is unknown. For example: Given the two earlier attributes geographic distance between an agent's primary activities and an agent's age, we want to know to what extent each distance, or similarity, between the attribute values of any two agents contributes to the similarity $d_{i,j}$ of the optimized activity chains of those two agents i and j . If the similarity of two agents' distances between their primary activities were twice as important than the similarity of their ages the distance metric would look like $\Delta = \{\delta_{distance_primacts}; \delta_{age}\} = \{2; 1\}$. If the similarities of the attributes were equally important the distance metric would look like $\Delta = \{\delta_{distance_primacts}; \delta_{age}\} = \{1; 1\}$, and so forth.

Solving the distance metric problem, we have developed a "reverse clustering" approach. The solution algorithm forms two sub-groups of agents. Group 1 is a group of optimized schedules (e.g., 100 agents). Group 2 is a group of non-optimized schedules (e.g., 500 agents). Both groups may be chosen randomly from among the overall set of agents. After having individually optimized the agents of group 1 through PlanomatX the algorithm assigns the non-optimized test agents with optimized schedules according to an arbitrary distance metric, e.g., $\Delta_t = \{\delta_1; \dots; \delta_k; \dots; \delta_n\} = \{1; \dots; 1; \dots; 1\}$. The new schedules of the agents will sum up to a certain utility $U_t = U_{\{1; \dots; 1; \dots; 1\}}$. The algorithm repeats the test assignment for x times. Each time, a set of varying distance metrics Δ_t is checked for the utility U_t they would lead to. The variation is done through increasing/decreasing each coefficient of the current metric by an offset value (e.g., 0.5). The alternative for which the utility becomes highest is chosen as the base metric for the following variation of coefficients. The variation undoing the previous

variation is forbidden. Finally, the algorithm selects the distance metric Δ_s for which the utility of the test agents' schedules is highest across all x times of the test assignment, i.e. $U_s > U_t$, $\forall t \neq s$.

6.3 Application tests and results

As for PlanomatX and the base test, we applied the schedule recycling concept to the greater Zurich scenario. The following paragraphs compare the schedule recycling results with the PlanomatX results.

6.3.1 Activity type constraints, discrete agent attributes and distance metric of continuous agent attributes

The check of activity type constraints always ensures that an agent may only be assigned with a schedule whose activity chain contains all primary activity types of the agent. Beyond that, we have defined the following constraints for an agent's eligibility of further activities:

- Children below 6 years may only be assigned with activity types of their initial schedule.
- Children between 6 and 17 years may be assigned with all available activity types but `education_kindergarten` and `education_higher`. Further, they may not be assigned with the work types except from if they hold them in their initial schedule.
- Adults from 18 years may be assigned with all available activity types except for `education_kindergarten`, `education_primary` and `education_secondary`.

We have omitted the check of discrete agent attributes given that currently no such (data on) discrete agent attributes exist that would prevent from recycling a schedule for any other agent. For the distance calculation between agents, we have selected the continuous attributes of

- agents' distances between their primary activities,
- agents' ages, and
- agents' home coordinates.

Based on these inputs, step C of the schedule recycling algorithm determined the distance metric $\Delta_s = \{\delta_{distance_primacts}; \delta_{age}; \delta_{home_coordinates}\} = \{2.5; 1.0; 1.5\}$.

6.3.2 Utility scores

The utility curves of PlanomatX and schedule recycling nearly overlap (figure 4). PlanomatX reaches a score level of 166 EUR, schedule recycling reaches a score level of 164 EUR. The difference of less than 2% is an excellent result. It demonstrates that our simple distance metric already generates good utility scores.

6.3.3 Optimized schedules

Columns 2 and 3 of table 1 contrast the activity chains statistics of the two tests. A first emphasis should be placed on the overall number of different activity chains: While PlanomatX has generated 652 different chains the schedule recycling produces only 457 different chains. This is reasonable since, in the schedule recycling mode, there is much less opportunity to produce activity chains than in the pure PlanomatX mode. It is in line with this finding that the 20 most frequent activity chains cover 160,785 agents (93.1% of all agents in the scenario). When we look at the frequency rankings of the activity chains we can observe that the 17 most frequent PlanomatX activity chains are within the 20 most frequent schedule recycling chains. The average length of the activity chains is similar being 5.38 activities for PlanomatX and 5.30 activities for the schedule recycling.

For sample agent 1000914, the PlanomatX and the schedule recycling activity chains are identical which is an excellent result (figure 5). The activity timings are very similar and, in both schedules, the agent uses the car throughout the day. The utility scores (169.47 EUR and 169.94 EUR) differ only in the decimal places.

6.3.4 Runtime

The runtime of the schedule recycling test was 45:04 hours. Each MATSim iteration took about 54 minutes splitting into 15 minutes of traffic flow simulation and 39 minutes of replanning runtime (0.134 seconds per agent). This is a massive runtime reduction compared to the PlanomatX test. The overall runtime has reduced by 84%, the replanning runtime has reduced by 88%. Moreover, the runtime has become more proportional with the initial base test's runtime: the additional activity chain dimension now leads to a replanning runtime increase of only 415%, as opposed to the 4,100% increase in the PlanomatX test.

6.4 Discussion

Schedule recycling is a concept to re-use schedules of optimized agents for other non-optimized agents. It avoids running PlanomatX for each agent individually and reduces the runtime by more than 80% compared to the pure application of PlanomatX. The diversification of activity chains after the schedule recycling is lower than after the PlanomatX optimization. However, the average executed utility score of the recycled schedules is only less than 2% worse than the score of the individually optimized PlanomatX schedules. This slight quality loss seems affordable, particularly given that only schedule recycling allows to handle large-scale scenarios such as the Greater Zurich Scenario at reasonable runtimes.

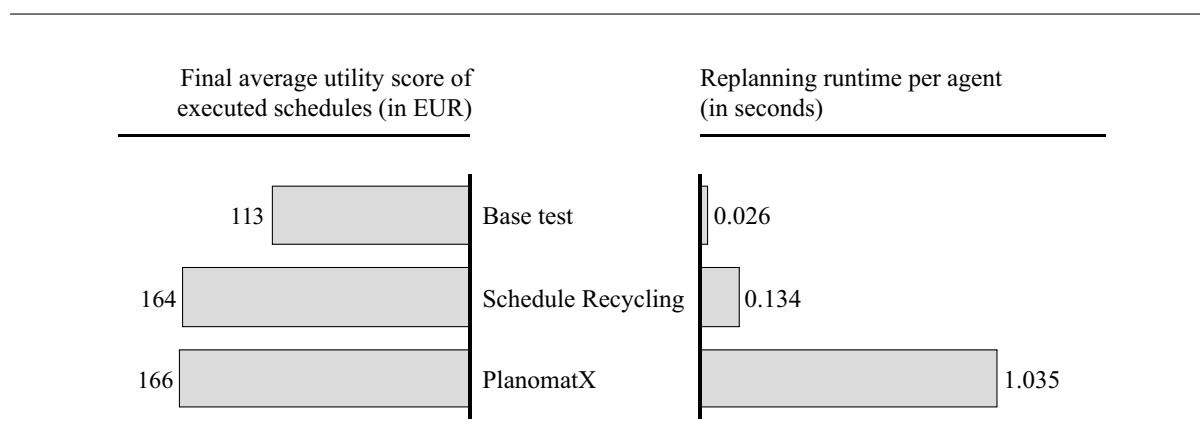
7 Conclusion and outlook

This paper presented the new scheduling algorithm PlanomatX as well as a new concept of schedule recycling. Both are part of the agent-based microsimulation MATSim. PlanomatX generates comprehensively optimized all-day schedules, i.e. optimal combinations of a schedule's activity chain (number, type and sequence of activities), activity timings, and the location, mode and route choices. It is based upon Tabu Search. Tests on the large-scale Greater Zurich Scenario with more than 170,000 agents have shown that PlanomatX successfully optimizes agents' schedules. In comparison with the joint application of MATSim's existing re-planning strategies (base test), PlanomatX has reached significantly higher utility scores. This has been possible since the structure of the activity chains has been a dimension of MATSim's co-evolutionary learning process, too. However, PlanomatX requires disproportional runtimes. Aiming for reduced runtimes, we have developed a concept of schedule recycling that avoids running PlanomatX for each agent individually. Schedule recycling rather re-uses schedules of optimized agents for other non-optimized agents. It significantly reduces simulation runtimes at affordable quality losses. Figure 6 summarizes utility scores and runtimes of PlanomatX, schedule recycling and reference base test.

We enhanced MATSim's existing utility function for the performance of activities. The existing function would have led, in combination with PlanomatX, to schedules with a lot of very short activities due to the decreasing marginal utility of its log-form. We have replaced the function by an asymmetric S-shaped curve with an inflection point, as presented by Joh (2004). The new function can cope with a flexible number of activities in the schedule as it formulates an optimal activity duration by its functional form.

Our future research will concentrate on improving the schedule recycling's distance metric as well as the utility function. Considering more agents attributes will increase the explanatory power of the distance metric. Furthermore, a more sophisticated search algorithm may enhance

Figure 6: Overview of utility score and runtime performances of the different tests.



the definition of the distance metric. With regard to the utility function, we will

- disaggregate the existing attributes of the function (e.g., more activity types, differentiation of the travel disutility by mode, etc.),
- incorporate new attributes (e.g., monetary cost of travelling), and
- empirically estimate the parameters of the utility function.

The results underpin that the last point will be central. The current utility function parameters together with the flexible activity chain dimension lead to a quite low variance in activity chains. The empirical estimation will better calibrate the parameters and set the results up for a comparison with real traffic counts.

Finally, an extension of PlanomatX and schedule recycling to optimize not only individual agents but entire households will be tackled in line with MATSim's overall ability to handle households.

References

- Arentze, T. A. and H. J. P. Timmermans (2004) ALBATROSS - a learning-based transportation oriented simulation system, *Transportation Research Part B: Methodological*, **38** (7) 613–633.
- Axhausen, K. W. (1988) Eine ereignisorientierte Simulation von Aktivitätenketten zur Parkstandswahl, Ph.D. Thesis, University of Karlsruhe, Karlsruhe.
- Balmer, M. (2007) Travel demand modeling for multi-agent traffic simulations: Algorithms and systems, Ph.D. Thesis, ETH Zurich, Zurich, May 2007.

- Balmer, M., K. Meister, M. Rieser, K. Nagel and K. W. Axhausen (2008a) Agent-based simulation of travel demand: Structure and computational performance of MATSim-T, paper presented at *Innovations in Travel Modeling (ITM'08)*, Portland, June 2008, <http://www.trb-forecasting.org/innovationsConference2008.html>.
- Balmer, M., K. Meister, M. Rieser, K. Nagel and K. W. Axhausen (2008b) Agent-based simulation of travel demand: Structure and computational performance of MATSim-T, *Working Paper*, **504**, IVT, ETH Zurich, Zurich, June 2008, <http://www.ivt.ethz.ch/vpl/publications/reports/ab504.pdf>.
- Balmer, M., M. Rieser, K. Meister, D. Charypar, N. Lefebvre and K. Nagel (2009) MATSim-T: Architecture and simulation times, in A. L. C. Bazzan and F. Klügl (eds.) *Multi-Agent Systems for Traffic and Transportation Engineering*, 57–78, Information Science Reference, Hershey.
- Bhat, C. R. (1997) Work travel mode choice and number of non-work commute stops, *Transportation Research Part B: Methodological*, **31** (1) 41–54.
- Bhat, C. R. (1998) A model of post home-arrival activity participation behavior, *Transportation Research Part B: Methodological*, **32** (6) 387–400.
- Bhat, C. R., J. Y. Guo, S. Srinivasan and A. Sivakumar (2004) A comprehensive econometric microsimulator for daily activity-travel patterns, *Transportation Research Record*, **1894**, 57–66.
- Bhat, C. R. and R. M. Pendyala (2005) Modeling intra-household interactions and group decision making, *Transportation*, **32** (5) 443–448.
- Bowman, J. L. and M. E. Ben-Akiva (1996) Activity-based travel forecasting, paper presented at *the Activity-Based Travel Forecasting Conference*, New Orleans, June 1996, <http://tmip.fhwa.dot.gov/clearinghouse/docs/abtf>.
- Bowman, J. L. and M. E. Ben-Akiva (2001) Activity-based disaggregate travel demand model system with activity schedules, *Transportation Research Part A: Policy and Practice*, **35** (1) 1–28.
- Bradley, M. A. and P. Vovsha (2005) A model for joint choice of daily activity pattern types of household members, *Transportation*, **32** (5) 545–571.
- Charypar, D. and K. Nagel (2005) Generating complete all-day activity plans with genetic algorithms, *Transportation*, **32** (4) 369–397.
- Cirillo, C. and K. W. Axhausen (Forthcoming) Dynamic model of activity type choice and scheduling, *Transportation*.

- Ettema, D. F. and H. J. P. Timmermans (eds.) (1997) *Activity-Based Approaches to Travel Analysis*, Pergamon, Oxford.
- Feil, M. (Forthcoming) Choosing the daily schedule: Expanding activity-based travel demand modelling, Ph.D. Thesis, ETH Zurich.
- Gliebe, J. P. and F. S. Koppelman (2005) Modeling household activity-travel interactions as parallel constrained choices, *Transportation*, **32** (5) 449–471.
- Glover, F. (1989) Tabu search - part 1, *ORSA Journal of Computing*, **1** (3) 190–206.
- Golledge, R. G., M.-P. Kwan and T. Gärling (1994) Computational-process modelling of household travel decisions using a geographical information system, *Papers of the Regional Science Association*, **73** (2) 99–117.
- Goulias, K. G. (ed.) (2002) *Transportation Systems Planning: Methods and Applications*, CRC Press, New York.
- Habib, K. M. N. and E. J. Miller (2008) Modelling daily activity program generation considering within-day and day-to-day dynamics in activity-travel behaviour, *Transportation*, **35** (4) 467–484.
- Hasan, M., T. AlKhamis and J. Ali (2000) A comparison between simulated annealing, genetic algorithm and tabu search methods for the unconstrained quadratic pseudo-boolean function, *Computers and Industrial Engineering*, **38** (3) 323–340.
- Hillier, F. S. and G. J. Lieberman (eds.) (2005) *Introduction to Operations Research*, vol. 8, McGraw-Hill, New York.
- Hobeika, A. (2005) TRANSIMS Overview, *Technical Report*, Virginia Polytechnic University, Virginia, July 2005, http://tmip.fhwa.dot.gov/transims/transims_fundamentals/ch1.pdf.
- Horni, A., D. M. Scott, M. Balmer and K. W. Axhausen (2008) Location choice modeling for shopping and leisure activities with MATSim: Combining micro-simulation and time geography, paper presented at *the 8th Swiss Transport Research Conference*, Ascona, October 2008.
- Illenberger, J., and K. Nagel (2007) Enhancing MATSim with capabilities of within-day re-planning, paper presented at *Intelligent Transportation Systems Conference (ITSC)*, Seattle, September 2007.
- Joh, C.-H. (2004) Measuring and predicting adaptation in multidimensional activity-travel patterns, Ph.D. Thesis, Technical University Eindhoven, Eindhoven.

- Kitamura, R. (1996) Applications of models of activity behavior for activity based demand forecasting, paper presented at *the Activity-Based Travel Forecasting Conference*, New Orleans, June 1996, <http://tmip.fhwa.dot.gov/clearinghouse/docs/abt.f>.
- Kitamura, R. and S. Fujii (1998) Two computational process models of activity-travel behavior, in T. Gärling, T. Laitila and K. Westin (eds.) *Theoretical Foundations of Travel Choice Modeling*, 251–279, Pergamon, Oxford.
- Lefebvre, N. and M. Balmer (2007) Fast shortest path computation in time-dependent traffic networks, paper presented at *the 7th Swiss Transport Research Conference*, Ascona, September 2007, http://www.strc.ch/conferences/2007/2007_lefebvre.pdf.
- MATSim-T (2008) Multi Agent Transportation Simulation Toolkit, webpage, <http://www.matsim.org>.
- Meister, K., M. Frick and K. W. Axhausen (2005) A GA-based household scheduler, *Transportation*, **32** (5) 473–494.
- Meister, K., M. Rieser, F. Ciari, A. Horni, M. Balmer and K. W. Axhausen (2009) Anwendung eines agentenbasierten Modells der Verkehrsnachfrage auf die Schweiz, *Straßenverkehrstechnik*, **53** (5) 269–280.
- Meyer, M. D. and E. J. Miller (2001) *Urban Transportation Planning*, 2. edn., McGraw-Hill, Singapore.
- Oi, W. Y. and P. W. Shuldiner (1962) *An Analysis of Urban Travel Demands*, Northwestern University Press, Evanston.
- Pendyala, R. M., R. Kitamura, A. Kikuchi, T. Yamamoto and S. Fujii (2005) Florida activity mobility simulator: Overview and preliminary validation results, *Transportation Research Record*, **1921**, 123–130.
- Pendyala, R. M., R. Kitamura, P. D. Reddy and Q. C. Chen (1995) Activity-based modeling system for travel demand forecasting, *Final Report*, Metropolitan Washington Council of Governments and Department for Transport, RDC Inc., San Francisco, September 1995, <http://tmip.fhwa.dot.gov/clearinghouse/docs/Amos/amos.pdf>.
- Rahoual, M. and R. Saad (2006) Solving timetabling problems by hybridizing genetic algorithms and tabu search, paper presented at *6th Interational Conference on the Practice and Theory of Automated Timetabling*, Brno, August 2006, http://patat06.muni.cz/doc/PATAT_2006_Proceedings.pdf.
- Recker, W. W., M. G. McNally and G. S. Root (1986a) A model of complex travel behavior: Part I - theoretical development, *Transportation Research Part A: Policy and Practice*, **20** (4) 307–318.

- Recker, W. W., M. G. McNally and G. S. Root (1986b) A model of complex travel behavior: Part II - an operational model, *Transportation Research Part A: Policy and Practice*, **20** (4) 319–330.
- Srinivasan, S. and S. R. Athuru (2005) Analysis of within-household effects and between-household differences in maintenance activity allocation, *Transportation*, **32** (5) 495–521.
- Srinivasan, S. and C. R. Bhat (2005) Modeling household interactions in daily in-home and out-of-home maintenance activity participation, *Transportation*, **32** (5) 523–544.
- Timmermans, H. J. P. (2001) Models of activity scheduling behavior, *Stadt Region Land*, **71**, 33–47.
- Timmermans, H. J. P. (2003) Recent and expected direction in modelling transportation, paper presented at *the 22nd Southern African Transport Conference*, Pretoria, July 2003, [https://www.up.ac.za/dspace/bitstream/2263/6578/1/Timmermans_Recent\(2003\).pdf](https://www.up.ac.za/dspace/bitstream/2263/6578/1/Timmermans_Recent(2003).pdf).
- TRANSIMS (2009) TRansportation ANalysis and SIMulation System, webpage, May 2009, <http://transims-opensource.org/>.
- Vovsha, P., E. Petersen and R. Donnelly (2002) Microsimulation in travel demand modeling: Lessons learned from the New York best practice model, *Transportation Research Record*, **1805**, 68–77.