Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 00 (2013)
Procedia Computer Science
www.elsevier.com/locate/procedia

The 2nd International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications (ABMTRANS'13)

Agent-based Modelling and Simulation of Air Transport Passenger Demand

D. Grether, S. Fürbas, K. Nagel

Transport Systems Planning and Transport Telematics, Berlin Institute of Technology, Salzufer 17-19, D-10587 Berlin, Germany, Tel: +49-30-314 21383, Fax: +49-30-314 26269, E-mail: grether@vsp.tu-berlin.de,

Abstract

With diminishing travel time differences between mid-distance transport modes analysis and forecasting needs to consider time-dependent inter-modal reactions. This paper aims at modelling such reactions using a multi-agent simulation approach. A simulation model for air transport technology is used, that represents details of air traffic microscopically and is fast enough to enable an iterative simulation-based passenger-trip assignment. Passengers are modelled as agents that may have individual attributes. In this paper, passengers differ by choice of departure time and chosen connection. Results for flights within Germany are presented, that are based on real world data. Overall, the approach appears to be suited to analyze and forecast mid-distance transport.

Keywords:
air transport planning, passenger-trip assignment, agent-based simulation

1. Introduction

In Italy, recently a private company started providing 2.5 h non-stop train rides between Milano and Rome.[1] From Paris, nearly all major French cities can be reached by high-speed train in 2–4 h trips.[2] For the journey Berlin–Frankfurt in Germany, a 4h non-stop rail connection is provided.[3] Many airlines provide flights between all these destinations that take between 1 and 2 h. When comparing travel times, the additional access time to the airport or railway railway station needs to be included. Overall travel times are often not that different between middle range rail on the one side, and air transportation on the other.

Following recent forecasts, in 2030 13 major EU airports will operate at full capacity at least eight hours a day [1]. Legal opening hour constraints limit operations to a certain time frame. Yet even increasing opening hours for airports may not resolve capacity bottlenecks since it may not be possible to move enough demand away from the peak hours.

[1] `http://www.italotreno.it` last access 19.12.2012
[2] `www.tgv-europe.com/en/`, last access 11.09.2012
[3] `www.bahn.com`, last access 11.09.2012

In contrast, railway stations are normally not as much exposed to restrictions of opening hours due to noise protection as airports are. Also, in comparison with airports, railway stations mostly feature a more central geospatial location in urban areas. Slightly longer travel times can be compensated for by shorter access times and longer opening hours. Passenger demand and technology supply for middle distance railway or air transportation may interact and are time dependent over a day or even a longer period.

In order to provide more capacity, railway or air transport networks may be target of planned extensions. New infrastructure is often accompanied by new emissions of noise and pollutants and is thus subject to lenghty planning, negotiation, and high private and public costs [2]. However, improvements on infrastructure may improve quality of journeys or offer even new possibilities of transportation. Identification and appraisal of these disadvantages and benefits is one of the key subjects in infrastructure planning.

Mutual reactions on several scales may arise if one or several transport measures cause disbenefits and advantages for certain user groups or individuals. For each transport system user, changes in price, travel times, schedule, or available transport modes have different impacts, which depend on planned activities, available budget and geospatial location. In order to analyze, forecast, and assess changes in air transport infrastructure and service, this paper employs a simulation and forecasting approach for individual passenger reactions, using concepts of multi-agent simulation for urban transport forecasting. The technology of air transport networks, i.e. airports and aircraft, are microscopically modelled the same way as buses or street cars. Details concerning modelling and simulation of air transport technology are presented in [3]. This paper focuses on passengers that are represented microscopically as multi-agent demand for air transportation. For modelling and simulation standard theory and tools available around the MATSim software toolkit are used. While the modelling is kept simple, the paper discusses several possibilities how it could be refined in order to provide answers to more specific questions.

The paper is organized as follows. Sec. 2 presents an overview of the methodology and the simulation framework. Then, Sec. 3 explains how passengers are included in the model while Sec. 4 describes simulation setup and results from several simulation runs. The paper ends with a discussion and conclusion.

2. Multi-Agent Transport Simulation

The simulation approach used in this paper is based on the software tool MATSim[4]. The next paragraphs provide an overview of the simulation approach and highlight the most important details used in this work. For more detailed information on technical aspects, please see [4] or [5]. For a detailed discussion of methodology, see, e.g., [6]. Regarding economic concepts used in the simulation approach, see, e.g. [7, 8].

Simulation overview. In MATSim, each traveler of the real system is modeled as an individual virtual person. The approach consists of an iterative loop that has the following important steps:

1. **Plans generation**: All virtual persons independently generate daily *plans* that encode, among other things, their desired activities during a typical day as well as the transportation mode. Virtual persons typically have more than one plan ("plan database").
2. **Traffic flow simulation**: All selected plans are simultaneously executed in a simulation of the physical system (often called "network loading").
3. **Scoring**: All executed plans are scored by an *utility function* which can be personalized for every individual.
4. **Learning**: At the beginning of every iteration, some virtual persons obtain new plans by modifying copies of existing plans. This is done by several *modules* that correspond to the choice dimensions available, e.g. time choice, route choice, and mode choice. In this paper, time and route choice will be used. Virtual persons choose between their plans according to a Random Utility Model (RUM).

[4]Multi-Agent Transport Simulation, see `www.matsim.org`.

The repetition of the iteration cycle coupled with the plan database enables the virtual persons to improve (learn) their plans over many iterations. This is why it is also called **learning mechanism** which is described in more detail by [5]. The iteration cycle continues until the system has reached a relaxed state. At this point, there is no quantitative measure of when the system is "relaxed"; we just allow the cycle to continue until the outcome is stable. In the steady state, the model is equivalent to the standard multinomial logit model

$$p_j = \frac{e^{\mu \cdot V_j}}{\sum_i e^{\mu \cdot V_i}} , \tag{1}$$

where p_j is the probability for plan j to be selected and μ is a sensitivity parameter, set to 2 for the simulations in this paper. In consequence, V corresponds to the systematic component of utility in Random Utility Models (RUM) e.g. [9, 10], where utility is defined as $U = V + \epsilon$. In RUM, the ϵ is called random component of utility. In the steady state and assuming a Gumbel distribution for ϵ, the choice model used in this paper is thus equivalent to the standard multinomial logit model.

Scoring. In order to measure the quality of a plan after execution and to compare plans, it is necessary to assign a quantitative score to the performance of each plan. For this purpose the utility function of the virtual persons is used. The total utility of a plan is computed as the sum of individual contributions:

$$V_{total} = \sum_{i=1}^{n} V_{perf,i} + \sum_{j=1}^{n} V_{tr,j} , \tag{2}$$

where V_{total} is the total utility for a given plan; n is the number of activities, which equals the number of trips (the first and the last activity are counted as one); $V_{perf,i}$ is the (positive) utility earned for performing activity i; and $V_{tr,j}$ is the (usually negative) utility earned for travelling during trip j. For calculation of $V_{perf,i}$ a logarithmic form is used

$$V_{perf,i}(t_{perf,i}) = \beta_{perf} \cdot t_{*,i} \cdot \ln\left(\frac{t_{perf,i}}{t_{0,i}}\right) \tag{3}$$

where t_{perf} is the actual performed duration of the activity, t_* is the "typical" duration of an activity, and β_{perf} is the marginal utility of an activity at its typical duration. β_{perf} is the same for all activities, since in equilibrium all activities at their typical duration need to have the same marginal utility. In this paper a β_{perf} of 6 utils is used. The (dis)utility of traveling is linear in travel time, i.e. $V_{tr,j}(t_{tr,j}) = \beta_{tr} \cdot t_{tr}$. In this work, β_{tr} is set to -6 for all virtual persons.

Further details on the default MATSim utility function can be found in [11] while one of the most recent discussions of this utility based approach is in [8].

3. Modelling of Passenger Demand for Air Transport

With the results from [3] an air transport technology model is available. This section shows how a passenger demand for air transport can be modelled with the multi-agent approach.

There are many different ways in which passenger demand for transport systems can be generated e.g. [12]. One option is to start with origin-destination (O-D) flows between geographical regions. In a European context, possible data-sources include OAG Aviation[5] and eurostat[6]. They provide data about passengers; O-D flows, however, are not provided. Data-sources geographically limited to Germany as "Der Flughafenverband"[7] or [13] do not come with O-D data, neither. The latter may have O-D relationships available in an upcoming version. The German Institute of Air Transport and Airport Research (DLR) provides monthly statistics containing O-D flows,[8] but the pdf format provided is not suited for machine

[5] `www.oagaviation.com`, last access 08.08.2012
[6] `ec.europa.eu/eurostat`, last access 10.09.2012
[7] `www.adv.aero`, last access 10.09.2012
[8] `http://www.dlr.de/fw/en/desktopdefault.aspx/tabid-2961/9753_read-19683/`, last access 10.09.2012

reading, and data is only available up to 09/2010. DESTATIS[9] provides O-D data by airport for German air traffic in a machine readable format. Data is available for whole years or a specific month. DESTATIS data is thus used in the following to create an agent based air transport demand for Germany.

The passenger demand is based on the DESTATIS data for 09-2011 in order to be consistent with [3]. DESTATIS provides data in two different representations (data sets 2.2.1 and 2.2.2). The number of O-D trips between airports is captured in two different ways. For all pairs of airports, the number of direct trips between the airports is given in the data set 2.2.1. Furthermore, the second data set 2.2.2 contains O-D pairs that do not include the first transfer, but provide the second, and possibly final destination. E.g. one person flying from Hamburg (HAM) via Frankfurt (FRA) to Munich (MUC) is contained in the data as one O-D pair: HAM $\rightarrow$ MUC. If a flight starts at Paris (CDG) going via FRA via MUC to HAM it is not clearly stated how the flight is represented in the data. It might be counted as CDG $\rightarrow$ HAM or FRA $\rightarrow$ HAM O-D relation. It is, however, unlikely that passengers will have two *transfer* stops within Germany. Thus an origin or destination abroad may not be the original or final destination, but at least all passenger movements that touch Germany along their itinery are probably included in an unequivocal way.

The second data set (2.2.2) is used to create the virtual persons for the passenger demand. For each O-D pair the number of trips is scaled from monthly to daily values by a division by 30. If the origin or destination airport is available in the simulation model for air transport technology, for each O-D pair and trip a virtual person is created, otherwise the trip is neglected. The resulting synthetic population contains 64006 virtual persons, 1103 trips from the original data are neglected. Each virtual person performs two activities, one at the origin and the other at the destination airport. Both activities are of same type, thus time spend performing both activities is accumulated before it is evaluated by the utility function according to Eq. (3). A "typical duration" (t_*) of 24 h is set for this activity type. In between the two activities a flight leg is scheduled, connecting origin and destination. As is common, the demand does not specify if a direct flight from O to D is chosen or the virtual person is on a route containing one or more transfers. The time virtual persons arrive at the origin airport and start waiting for a connection is drawn randomly from a uniform distribution in 04:00 to 18:00, UTC. This reflects estimated typical opening hours of airports in Europe.

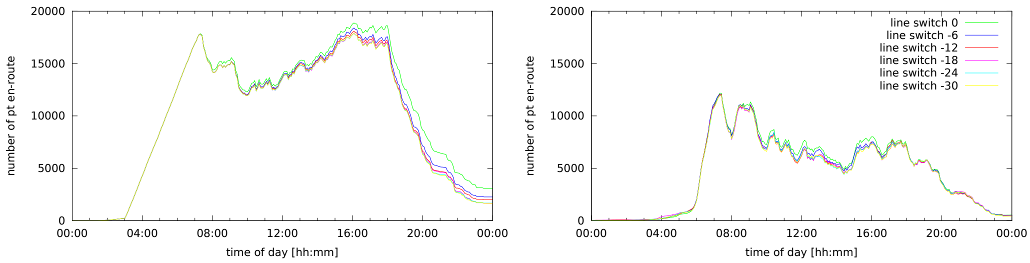
4. Simulation Setup & Results

The synthetic population is used as input for the simulation. As scenario the European flight model with no delays and no effective runway capacity restrictions from [3] is used. The assignment of concrete flights to the desired O-D connection, i.e. the passenger routing, is done by the default public transit routing module of MATSim [14]. The routing basically looks for a least cost path in terms of travel time. The network used for routing is constructed from the information contained in the transit schedule. In order to penalize transfers, the routing assumes an additional cost of $c_{lineswitch}$ for each transfer. The same parameter is also considered by the scoring function, i.e. a (dis-)utility of $-c_{lineswitch}$ is added to the score of the agent for each transfer. The simulation is run several times using different values of the $c_{lineswitch}$ parameter, i.e. $c_{lineswitch} \in \{0, -6, -12, -18, -24, -30\}$[utils/transfer].

The simulation is run for 600 iterations. In each iteration, 10 % of the virtual persons may shift their departure time randomly within a 2 h interval. The amount of shift is drawn from a uniform distribution. Another 10 % may seek a different route, i.e. a different connection between origin and destination. Each passenger chooses out of a set of 5 plans using a multinomial logit model, see [7] for details. The outcome is stable after 500 iterations, thus departure time choice and routing are switched off. For another 100 iterations only the logit model is used by the passengers to select a plan. Empirically, fixing the choice set for the last 100 iterations reduces the noise of learning and eases analysis and intepretation of results. All other parameters used for simulation are the "default" values of the MATSim framework. For a detailed discussion, see, e.g., [7, 6].

One iteration takes around 10 min. on an Intel Xeon Processor (2.67 GHz) using one core for the execution of mobility simulation and two cores for the replanning modules.

[9]`destatis.de`, Fachserie 8 Reihe 6, last access 10.09.2012



(a) After 0th iteration (b) After 600th iteration

Fig. 1: Travelers en-route, i.e. waiting for a flight or travelling by plane, as a function of the time-of-day

First, in order to show the effects of routing, the result after the zeroth simulated iteration is presented. Each virtual person gets a connection assigned based on a generalized cost routing for the connection and the preset departure time. Fig 1a shows the number of travellers en-route, i.e. waiting for a flight or travelling by plane, as a function of the time-of-day. Some passengers are still waiting for a flight at midnight. As only one day of operation is simulated, these passengers are stuck and will not reach their destination. The number of stuck passengers is decreasing with the increasing disutility of line switch.

The output after 600 iterations is depicted in Fig. 1b. The shape of all curves is different from the shape of the 0th iteration. One can identify two morning and two evening peaks. Some passengers still get stuck at the end of the day, but fewer than in the 0th iteration. In addition, the differences between the curves for the $c_{lineswitch}$ parameter are diminishing.

In order to study the influence of the $c_{lineswitch}$ parameter, the simulation results are compared with the leg data. Recall that the synthetic population is generated based on O-D pairs that may contain transfers ($od_{transfers}$), while other data directly counts the number of passengers on actual direct flights (od_{direct}). The latter is used to evaluate the accuracy of the model. For comparison, the number of passengers on direct flights is thus calculated for each O-D pair (sim_{direct}) from the simulation results.

Based on these datasets, the mean square error σ^2 is computed as

$$\sigma^2 = \frac{\sum_{i \in OD}(sim_{direct}(i) - od_{direct}(i))^2}{|OD|} .$$

The (unsigned) mean relative error for each O-D relation is calculated as

$$\text{mean rel error} = \frac{\sum_{i \in OD} |(sim_{direct}(i) - od_{direct}(i))| / od_{direct}(i)}{|OD|} .$$

Tab. 1 shows the results for these calculations. The first line contains the comparison of the two input data sets from DESTATIS, i.e. in the above formulas sim_{direct} is replaced by $od_{transfers}$. This serves as reference as it would assume that *all* demand is served by direct flights. All simulation runs explain the data better than that reference. The values for all simulation runs are then quite similar.

The last column of tab. 1 shows the number of passengers stuck at the end of day. Values for all parameter settings are around 980 passengers, i.e. around 1,5 % of the 64'006 simulated passengers.

5. Discussion

Overall, the results show that a microscopic, agent-based simulation of passenger demand for air transport is feasible. Most passengers are able to learn the constraints of air transport technology and arrive at their desired destination.

Some passengers fail to reach their destination; they get "stuck". As only trips within Germany are modelled, which are usually completed within a few hours without any requirement for an overnight stay

$c_{lineswitch}$	σ^2	σ	mean rel error	stuck
$od_{transfer} - od_{direct}$	6714	81	1.56	-
-0	5329	73	0.29	982
-6	5663	75	0.27	976
-12	5801	76	0.28	967
-18	5838	76	0.24	985
-24	5889	76	0.25	968
-30	5935	77	0.25	971

Table 1: Simulation results for different values of $c_{lineswitch}$

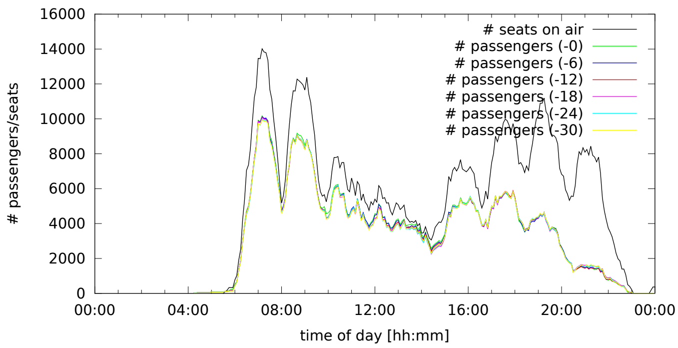


Fig. 2: Passengers and available seats over time within Germany

at an airport, this is considered unrealistic. Fig. 2 shows that this is not a consequence of a general lack of seats: at any time-of-day, there are more seats than demand. There are many reasons why stuck passengers can arise in such a situation. Further analysis of the simulation results leads to the following insights

- 459 passengers are stuck because there is no seat, *and* there is no other flight by the same airline later during the day to which they would be shifted otherwise.
- 523 passengers are stuck at an airport because there is no connection after their departure time between that airport and their destination airport.

These results are for the run with $c_{lineswitch} = 0$; the other runs are similar. It may, for example, be possible that passengers depart from an origin that only has one –early– connection to a hub per day, and the passengers' departure times are too late to reach that connection, and the random departure time mutation may not be able to find that connection for all passengers. Alternatively, it may be the case that passengers have a connection that works in theory, but they are "crowded out" by other passengers who arrive earlier at the gate. They would make it if either of them would take a different route. The current approach would not find such a solution, since passengers do not take into account the costs they impose on others, see [15] for an approach to take that into account. The real-world solution presumably would be to raise prices on congested seats until one or the other passenger re-routes. The present model does not (yet) include such a process.

An alternative approach to remove some of these shortcomings might be to use a router that generates a larger diversity of routes even for the same departure time. Such a router would be able to point a passenger to a route where seats are available without by itself knowing about seat availability. That approach would, however, not address the issue that some passengers might need to switch their path in order to allow *others* to obtain a feasible path.

6. Conclusion

This paper aims at improving mid-distance transport. It shows how an agent-based demand for air transport can be derived from O-D flows. Use of a computationally affordable simulation technique enables an iterative assignment to flights. Results are presented for a simulation of German national air transport demand. It is shown that knowledge gained from input data can be improved by the learning mechanism of the agents. Overall, results look promising, the paper discusses several possibilities of model refinement.

Alternative transport modes should be included into the modelling. Models for other modes as rail or car transportation are the subject of current work, following the same approach as the one presented here. The same software and solution procedure is used. In consequence, these models can be integrated into the approach presented in this paper. This might help to get a more detailed picture of middle distance traffic modelling.

Acknowledgements

For the model and simulation approach proposed in this paper several prototype studies have been conducted [16, 17]. These prototypes have been target to further improvement and development, the paper describes the overall result. Further acknowledgements to all providers of data used in this study. Without accessible data suited for machine processing this work would not have been possible.

References

[1] E. Commission, Airport policy in the European Union – addressing capacity and quality to promote growth, connectivity and sustainable mobility, COM(2011) 823 final (12 2011).
URL `http://ec.europa.eu/transport/air/airports/airports_en.htm`
[2] B. Bubalo, J. Daduna, Airport capacity and demand calculations by simulation–the case of Berlin-Brandenburg International Airport, NETNOMICS (2012) 1–21doi:10.1007/s11066-011-9065-6.
[3] D. Grether, S. Fürbas, K. Nagel, Agent-based modelling and simulation of air transport technology, in: The 2nd International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications, 2013.
[4] B. Raney, K. Nagel, An improved framework for large-scale multi-agent simulations of travel behaviour, in: P. Rietveld, B. Jourquin, K. Westin (Eds.), Towards better performing European Transportation Systems, Routledge, London, 2006, pp. 305–347.
[5] M. Balmer, B. Raney, K. Nagel, Adjustment of activity timing and duration in an agent-based traffic flow simulation, in: H. Timmermans (Ed.), Progress in activity-based analysis, Elsevier, Oxford, UK, 2005, pp. 91–114.
[6] K. Nagel, G. Flötteröd, Agent-based traffic assignment: Going from trips to behavioural travelers, in: R. Pendyala, C. Bhat (Eds.), Travel Behaviour Research in an Evolving World – Selected papers from the 12th international conference on travel behaviour research, International Association for Travel Behaviour Research, 2012, Ch. 12, pp. 261–294.
URL `http://www.lulu.com/shop/ram-m-pendyala-and-chandra-r-bhat/travel-behaviour-research-in-an-evolving-world/hardcover/product-18838195.html`
[7] K. Nagel, D. Grether, U. Beuck, Y. Chen, M. Rieser, K. Axhausen, Multi-agent transport simulations and economic evaluation, Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik) 228 (2+3) (2008) 173–194, special issue on agent based models for economic policy advice, edited by B. LeBaron and P. Winker.
[8] B. Kickhöfer, D. Grether, K. Nagel, Income-contingent user preferences in policy evaluation: application and discussion based on multi-agent transport simulations, Transportation 38 (2011) 849–870. doi:10.1007/s11116-011-9357-6.
[9] M. Ben-Akiva, S. R. Lerman, Discrete choice analysis, The MIT Press, Cambridge, MA, 1985.
[10] K. Train, Discrete choice methods with simulation, Cambridge University Press, 2003.
[11] D. Charypar, K. Nagel, Generating complete all-day activity plans with genetic algorithms, Transportation 32 (4) (2005) 369–397.
[12] M. Balmer, Travel demand modeling for multi-agent transport simulations: Algorithms and systems, Ph.D. thesis, Swiss Federal Institute of Technology (ETH) Zürich, Switzerland (2007).
[13] ITP/BVU, Prognose der deutschlandweiten Verkehrsverflechtungen 2025, Tech. rep. (2005).
URL `http://daten.clearingstelle-verkehr.de/220/`
[14] M. Rieser, Adding transit to an agent-based transportation simulation concepts and implementation, Ph.D. thesis, TU Berlin, also VSP WP 10-05, see www.vsp.tu-berlin.de/publications (2010).
[15] G. Lämmel, G. Flötteröd, Towards system optimum: Finding optimal routing strategies in time-tependent networks for large-scale evacuation problems, Vol. 5803 of LNCS (LNAI), Springer, Berlin Heidelberg, 2009, pp. 532–539. doi:10.1007/978-3-642-04617-9_67.
[16] S. Fürbas, Modellierung und Simulation des europäischen Linien–Luftverkehrs, Bachelor's thesis, TU Berlin, Institute for Land and Sea Transport Systems, Berlin, Germany (03 2011).
[17] C. M. Treczka, Modellierung und Simulation der Luftverkehrsnachfrage im deutschen Luftraum, Bachelor's thesis, TU Berlin, Institute for Land and Sea Transport Systems, Berlin, Germany (01 2012).