Towards a more realistic simulation of public transit: Generating transit schedules with vehicle circulations

Gero L. Marburger[a], Ihab Kaddoura[a,*], Kai Nagel[a]

[a]*Technische Universitt Berlin; Department of Transport Systems Planning and Telematics; Salzufer 17-19; 10587 Berlin; Germany*

Abstract

In this study, two approaches are presented to account for vehicle circulations when incorporating GTFS data into the agent-based simulation framework MATSim. The first approach directly builds on an existing data converter; the second approach integrates the external public transit planning tool LinTim into MATSim. Both proposed methodological approaches are successfully used to generate transit schedules with vehicle circulations. Simulation experiments are carried out for the Greater Berlin area and reveal an overall minor impact on the public transit performance from the transport users' perspective. The number of required public transit vehicles is significantly reduced compared to previous version of the GTFS data converter. Reducing the turnaround time leads to a further reduction in vehicles, however, provides reduced slack times to compensate for delays. The implementation of a disruption shows that the bunching phenomenon and delay propagation is now represented by the model. Thus, the proposed methodology allows for a more sophisticated investigation of transit schedules and delay management concepts.

Keywords: MATSim; agent-based simulation; public transit; vehicle circulation; transit schedule; GTFS; LinTim

1. Introduction

The agent-based transport simulation framework MATSim (Multi-Agent Transport Simulation, www.matsim.org) is a widely used open-source software for multi-modal and large-scale traffic systems. For the simulation of public transit (PT), typically GTFS (General Transit Feed Specification) data is used to generate the required MATSim transit schedule and transit vehicles files. Existing software to convert GTFS data into the MATSim format focuses on the correct representation of departure times and travel times from the users' perspective [6]. However, the existing approaches do not generate vehicle circulations (vehicle schedules, vehicle cycles), instead a new vehicle is generated for each transit route's departure. In this study, MATSim's toolchain is extended by a converter which automatically generates vehicle circulations. This yields a more realistic representation of the PT system and allows for additional investigations, e.g. delay propagation [4], the bus bunching phenomenon [5] and cost calculations for the required rolling stock. Neumann and Nagel [5] use MATSim to investigate bus bunching phenomenon and develop strategies to mitigate this problem. Müller-Hannemann and Schnee [4] study the effects of delay propagation and secondary delays in PT and how these delay effects can be predicted by means of a software-based approach.

In this study, two different software approaches are pursued. The first approach is a software extension which is directly integrated into MATSim's toolchain and extends an existing GTFS data converter (GTFS2MATSim).

---

* Corresponding author. Tel.: +49-30-314-78793 ; fax: +49-30-314-26269.
  E-mail address: kaddoura@vsp.tu-berlin.de

This approach does not depend on any external transport planning software. The second approach incorporates the open-source planning software LinTim (Lineplanning and Timetabling, https://www.lintim.net) into MATSim, in particular a LinTim tool for the optimization of vehicle circulations. The developed software and further documentation is publicly available: The extended GTFS2MATSim converter may be found here: https://github.com/matsim-org/GTFS2MATSim. The interface for LinTim and its documentation are located here: https://github.com/Marbventa/matsim-berlin/tree/5.5.x/src/main/java/org/matsim/run/linTimVehicleCirculation.

This paper is structured as follows. Sec. 2 describes the methodology including the technical implementations in MATSim. In Sec. 3, the proposed methodology is tested for the case study of the Greater Berlin area. Finally, Sec. 4 provides the conclusion and outlook.

2. Methodology

2.1. MATSim

In MATSim, travel demand is modeled as individual agents that execute, evaluate and adjust so-called daily plans. An agent's daily plan describes the agent's activities as well as trips to get from one activity to the next one. During their trips, agents use the modeled transport infrastructure. For PT trips, agents walk to the transit stop, wait at transit stops, get in and out of PT vehicles and change from one vehicle to another. Private and public vehicles move along the edges (roads) and nodes (intersections) of the modeled transport network. MATSim's transit schedule contains all departures of PT vehicles and their corresponding stops; the departures are assigned to transit lines and transit routes. Transit routes describe a sequence of stops to be served. For each stop, an off-set is defined which describes the departure time in relation to the departure time at the first stop of the TransitRoute. In addition to the stops, the TransitRoutes also specify the associated edges that are used to connect two stops. Transit stop facilities have coordinates and an edge; the stop is at the end of the edge. If a stop can be reached from more than one other stop, there must be multiple stop facilities, each with an edge for the previous stop. A majority of these multiple stop facilities are listed as parent-child stops and thus have the same stop area ID. It is possible to travel between these stops without an additional edge. A detailed description of MATSim's PT is provided in Rieser [7, 8].

For the generation of MATSim's PT data, in particular the transit schedule and the transit vehicles, there are two widely used and actively maintained converters: GTFS2MATSim (https://github.com/matsim-org/GTFS2MATSim) and PT2MATSim [6]. GTFS2MATSim is a converter primarily maintained and used by the TU Berlin team. PT2MATSim is mainly used by the ETH Zurich team and provides improved functionality in the context of locating bus lines on the road network. In this study, the proposed software extension is based on the existing GTFS2MATSim converter. The developed methodology may eventually also be incorporated in the PT2MATSim converter.

2.2. LinTim

LinTim is a unix-based open source software for planning and optimizing PT. Using PT infrastructure and the schedule as well as further configuration parameters, PT lines as well as the vehicles may be planned in an integrated way. In principle, planning and optimization are carried out in several different programs. These programs can be executed independently and represent individual planning steps for PT. In the vehicle scheduling planning step, vehicles are assigned to the trips to be served. The circulations are recorded in the vehicle schedule as a trip with a starting and ending stop. In addition, empty trips are created that connect the final stop of a line and the starting stop of the next line to be served [1, 2]. When using vehicle scheduling in the creation of cross-line vehicle circulation, a license of one of the supported commercial integer programming solvers is required. In this study, Gurobi is used because of its free academic license (https://www.gurobi.com/).

2.3. Proposed approaches to enable vehicle circulation

Extended GTFS2MATSim converter. The proposed methodology is implemented as an optional step in the existing GTFS2MATSim converter and is based on the following premises: (i) Vehicles will only serve departures. Other trips, such as empty trips or depositing on side tracks or in depots, are not considered. (ii) Vehicles only serve a single

transit line. There is no vehicle circulation between different transit lines. (iii) Vehicles may be parked at a stop for any duration. (iv) Turning loops are not considered, vehicles may turn around at any terminal stop. (v) Turnaround times are set by the user.

The developed approach aims to continuously schedule vehicles under consideration of the turnaround time. The approach first creates a list of all departures of a transit line sorted from earliest to latest departure time. If no transit vehicle is available, new vehicles are created to serve the predefined departures. In case a transit vehicle has completed a transit route and has arrived at the terminus stop plus the turnaround time has passed, the vehicle is considered available and may be assigned to the next departure. The assignment of start and end stop is made on the basis of the stop facility name. The name of the next start stop for a vehicle must be identical to the name of the end stop of the previous transit route. In case the first and last stop have different stop area IDs, connection edges are created between the end and start stop with a length of 50 meters, a speed of 10m/s and a sufficiently large capacity.

Coupling of MATSim and LinTim. The coupling of MATSim with the optimization program LinTim is based on the following premises: (i) Vehicles will only serve departures. Other trips, such as empty trips or depositing on side tracks or in depots, are not considered (same as above). (ii) PT vehicles of the same transport mode are considered as identical. Vehicles of the same mode may circulated freely between different transit lines. (iii) Vehicles may be parked at a stop for any duration (same as above). (iv) Turning loops are not considered, vehicles may turn around at any terminal stop (same as above). (v) Turnaround times are set by the user (same as above); here: in the configuration file of LinTim.

In contrast to the previously presented approach, the MATSim-LinTim coupling allows PT vehicles to be scheduled in cross-line circulations. The vehicle scheduling program uses the following three files to generate vehicle circulations: stops, edges and trips. For the integration of LinTim, all required files are generated directly from MATSim. Since LinTim does not distinguish vehicles by mode, the transit routes in MATSim are split into the respective modes. Stops are exported by their coordinates and not by their names. The existing links between stops in MATSim remain. Each departure in MATSim is represented as a separate line in LinTim. Only the start and end stops and the associated trip duration are exported. The previously existing vehicle IDs are replaced by new circulating vehicle IDs. New links are inserted based on the empty trips created by the vehicle scheduling. Only one edge per transport mode is added for each pair of nodes. The generated links have the same parameters as for the GTFS2MATSim program.

3. Application: Exemplary simulation for both software approaches

3.1. Simulation setup

For testing and comparing the developed approaches, the Open Berlin MATSim scenario is used. The Open Berlin MATSim scenario developed by Ziemke et al. [9] and maintained by the the Department of Transportation System Planning at TU Berlin contains a synthetic population for the Greater Berlin area and accounts for both the private and PT mode. The PT network incorporates the relevant infrastructure for bus and tramway as well as for suburban, underground and regional trains in the Berlin metropolitan area [9]. In this study, the simulation is run for a single iteration and a 10% sample of the population in the Greater Berlin area is used. The existing transit schedule is replaced by a transit schedule with circulating vehicles. For each approach, a transit schedule with a turnaround time of 120 minutes and 10 minutes is created. The 120 minutes are selected in order to keep the interference in the operation of the PT as low as possible. The turnaround time of 10 minutes is to model an optimal use of vehicles. The four newly generated configurations as well as the basic configuration of the Berlin scenario as reference scenario are simulated. The basis configuration is adjusted for this by setting the length of the links with length zero to one. This is necessary because otherwise the routing algorithm for circulating vehicles might run into a loop. To measure changes in PT operations, the following outputs of the simulations are evaluated: average score of agents (user benefits), travel time of agents, and number of PT vehicles. A small change in average score and travel time indicates a low impact of circulating vehicles on the simulation. A lower number of generated vehicles describes a more suitable algorithm for the generation of circulating vehicles.

3.2. Circulating vehicles with turnaround time of 120 minutes

As shown in Table 1, for a turnaround time of 120 minutes, the deviations of the average scores from the reference scenario are very small. The respective travel times hardly differ from the travel time in the reference scenario. The differences of the average scores as well as the average travel time of a PT leg are slightly higher in the LinTim simulation compared to the extended GTFS2MATSim approach, thus the circulations have a slightly stronger impact on the simulated transport users. The transit schedules generated with both vehicle circulation approaches require significantly fewer PT vehicles compared to the reference scenario where a transit vehicle is generated for each route's first departure.

Table 1. Comparison of the key indicators between the simulations of the reference scenario, extended GTFS2MATSim and LinTim integration with a turnaround time of 120 minutes. Other transport modes, e.g. ferry, are excluded from the vehicle circulation planning.

Key Indicator	Reference case	Extended GTFS2MATSim	LinTim Integration
Average score (*utility units*)	116.106794418118	116.106792817536	116.106786164826
Increase in avg. score (*utility units*)		0.000001600582	0.000008253292
Average travel duration of a PT leg (*sec*)	973.0144667669796	973.0151622846228	973.018024085267
Increase in avg. travel duration of a PT leg (*sec*)		0.0006955176432	0.0035573182874
Number of PT vehicles	72461	18838	14800
Number of bus vehicles	53082	15379	11299
Number of tramway vehicles	8964	2342	1493
Number of rail vehicles	10068	1043	1661
Number of ferries	347	74	347*

*Ferries are excluded from the vehicle circulation planning.

3.3. Circulating vehicles with a turnaround time of 10 minutes

As shown in Table 2, for a turnaround time of only 10 minutes, the average score has slightly decreased and the average PT travel times have slightly increased compared to the 120 minutes turnaround times. This indicates a

Table 2. Comparison of the key indicators between the simulations of the reference scenario, extended GTFS2MATSim and LinTim integration with a turnaround time of 10 minutes.

Key Indicator	Reference case	Extended GTFS2MATSim	LinTim integration
Average score (*utility units*)	116.106794418118	116.10678193653	116.10675711579
Increase in avg. score (*utility units*)		0.000012481588	0.000037302328
Average travel duration of a PT leg (*sec*)	973.0144667669796	973.0193235157908	973.030366978858
Increase in avg. travel duration of a PT leg (*sec*)		0.0048567488112	0.0159002118784
Number of PT vehicles	72461	11591	6887
Number of bus vehicles	53082	9235	5300
Number of tramway vehicles	8964	1511	583
Number of rail vehicles	10068	834	657
Number of ferries	347	11	347*

*Ferries are excluded from the vehicle circulation planning.

slightly stronger impact on the simulated transport users. Overall, the impact on transport users is still very small. At the same time, significantly fewer vehicles are needed to serve the schedule compared to the reference case. The transit schedule generated by the extended GTFS2MATSim approach includes more vehicles and affects the simulation less compared to the schedule generated using LinTim. The GTFS2MATSim approach only circulates vehicles within the same transit line. In contrast, the integration of LinTim into MATSim generates cross-line vehicle circulations.

3.4. Simulation of a disruption: one-hour delay with circulating vehicles and a turnaround time of 10 minutes

To test the implications of vehicle circulations a service disruption is implemented for subway line U9 using the Open Berlin scenario by Ziemke et al. [9]. Similar to Müller et al. [3] who examine the effects of a disruption on
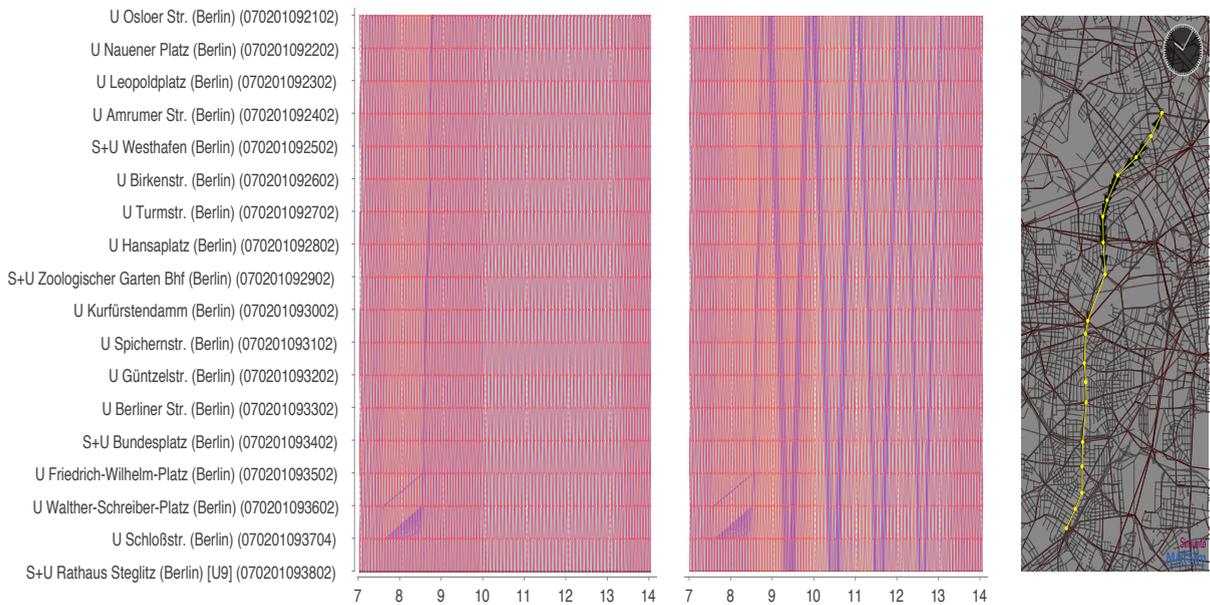
Fig. 1. Trajectories of subway line U9. **Left:** Simulation results in the reference case without vehicle circulations. **Center:** Simulation results for the extended GTFS2MATSim approach with vehicle circulations. Blue lines indicate simulated trajectories. Red indicates the transit schedule. **Right:** Snapshot from 10:05 am of the subway line U9 (yellow) in the extended GTFS2MATSim approach. PT Vehicles shown as black triangles.

passengers' route choices, a breakdown of subway line U9 is implemented during the time period from 7:30 am to 8:30 am. The simulation is run for a single iteration and user reactions as well as delay management policies by the operator are not accounted for. For the simulation experiments, a transit schedule is created which contains vehicle circulations with a turnaround time of 10 minutes. Figure 1 depicts the simulation outcome. On the left and in the center, PT vehicles' trajectories are shown with time of day (from 7:00 am to 2:00 pm) on the x-axis and transit stops of subway line U9 on the y-axis. Simulated trajectories are shown in blue and travel times defined in the schedule are shown in red. The figure on the left depicts the PT vehicle trajectories for a transit schedule without vehicle circulations (reference scenario). The central figure shows the trajectories for a schedule with vehicle circulations generated using the proposed GTFS2MATSim approach. The figure on the right provides the corresponding visualization of the simulated vehicles at 10:05 am. The subway line U9 is shown in yellow and PT vehicles are colored in black. The stop names of transit line U9 marked on the left are ordered from North (U Osloer Str.) to South (S+U Rathaus Steglitz) and correspond to the yellow dots shown on the right. The depicted trajectories show an interruption of the subway line before leaving the stop "U Walter-Schreiber-Platz" between 7:30 and 8:30 am. The initially delayed vehicles remain as a bunch, indicated by the accumulation of blue lines. Without vehicle circulations (figure on the left), there is no delay propagation. In contrast, in the case with vehicle circulations (figure in the center), PT vehicles remain delayed after arriving at the destination stop and departing for their next circulation. This remains to be the case until about 1:00 pm. The affected vehicles go through more than 6 vehicle circulations causing additional delays later during the day. Short delays can be mitigated by reducing the wait-time at stops. However, the boarding time of the passengers cause additional delays since passengers were unable to enter a vehicle for an hour causing an accumulation of passengers. As depicted in the figure on the right, vehicle bunching is still observed more than an hour after the disruption. Resulting vehicle trajectories for the transit schedule with vehicle circulations based on the LinTim approach show a similar pattern as in the extended GTFS2MATSim case (figure in the center).

Without vehicle circulation, the total delay amounts to 3,233 person-hours. With vehicle circulation the total delay increases by a factor of 3.17 in the GTFS2MATSim approach and by a factor of 3.06 in the LinTim integration approach. In the case without vehicle circulations 6,130 passengers experience an increase in travel time. For the simulation experiments with vehicle circulation this number is significantly larger. The number of delayed passengers increases by a factor of 4.36 in the GTFS2MATSim approach and by a factor of 4.50 in the LinTim approach.

4. Conclusion and outlook

In this study, two approaches are presented to account for vehicle circulations when incorporating GTFS data into the agent-based simulation framework MATSim. The first approach directly extends the existing GTFS2MATSim converter; the second approach integrates the external PT planning tool LinTim into MATSim. The extended GTFS2MATSim approach builds solely on the functionalities of MATSim and is thus directly integrated into MATSim's open-source software toolchain. The LinTim integration approach uses the open-source optimization program LinTim, but requires an integer programming solver.

Both proposed methodological approaches are successfully used to generate transit schedules with vehicle circulations. Simulation experiments are carried out for the Greater Berlin area and reveal an overall minor impact on the PT performance from the transport users' perspective. Scores and travel times remain at the approximately same level. By implementing vehicle circulations, the number of required PT vehicles is significantly reduced compared to previous version of MATSim's GTFS data converter. Reducing the turnaround time leads to a further reduction in vehicles, however, provides reduced slack times to compensate for delays. The comparison of both methodological approaches shows that the GTFS2MATSim approach generates more circulating vehicles and yields a slightly smaller impact on simulated transport users compared to the LinTim integration approach.

The implementation of a disruption shows that the bunching phenomenon and delay propagation is now represented by the model and may be investigated for different policies or transit schedule design concepts. The visual comparison between the trajectories of the subway line U9 for both approaches does not show a significant difference in terms of the bunching phenomenon or delay propagation.

In future studies, the LinTim integration in MATSim may be used as an external optimization tool for further PT planning steps. The GTFS2MATSim approach may be extended to account for vehicle circulations between different (selected) transit lines; with different turnaround times for different transit lines. Overall, the proposed methodology allows for a more sophisticated investigation of transit schedules and delay management concepts.

References

[1] S. Albert, J. Pätzold, A. Schiewe, P. Schiewe, and A. Schöbel. Documentation for lintim 2020.02, 2020. URL http://nbn-resolving.de/urn:nbn:de:hbz:386-kluedo-59130.

[2] M. Friedrich, M. Müller-Hannemann, R. Rückert, A. Schiewe, and A. Schöbel. Robustness as a third dimension for evaluating public transport plans. 2018. doi: 10.4230/OASICS.ATMOS.2018.4.

[3] S. A. Müller, G. Leich, and K. Nagel. The effect of unexpected disruptions and information times on public transport passengers: a simulation study. *Procedia Computer Science*, 170:745–750, 2020. doi: 10.1016/j.procs.2020.03.161.

[4] M. Müller-Hannemann and M. Schnee. Efficient timetable information in the presence of delays. In *Robust and Online Large-Scale Optimization*, pages 249–272. Springer Berlin Heidelberg, 2009. doi: 10.1007/978-3-642-05465-5_10.

[5] A. Neumann and K. Nagel. Avoiding bus bunching phenomena from spreading: A dynamic approach using a multi-agent simulation framework, 2010. URL https://svn.vsp.tu-berlin.de/repos/public-svn/publications/vspwp/2010/10-08/20100730_submitted.pdf.

[6] F. Poletti, P. Bösch, F. Ciari, and K. Axhausen. Public transit route mapping for large-scale multimodal networks. *ISPRS International Journal of Geo-Information*, 6(9):268, aug 2017. doi: 10.3390/ijgi6090268.

[7] M. Rieser. Adding transit to an agent-based transportation simulation: Concepts and implementation, 2010.

[8] M. Rieser. Modeling public transport with MATSim. In *The Multi-Agent Transport Simulation MATSim*, chapter Modeling Public Transport with MATSim, pages 105–110. Ubiquity Press, aug 2016. doi: 10.5334/baw.16.

[9] D. Ziemke, I. Kaddoura, and K. Nagel. The MATSim open berlin scenario: A multimodal agent-based transport simulation scenario based on synthetic demand modeling and open data. *Procedia Computer Science*, 151:870–877, 2019. doi: 10.1016/j.procs.2019.04.120.