Dette resumé er publiceret i det elektroniske tidsskrift
Artikler fra Trafikdage på Aalborg Universitet
(Proceedings from the Annual Transport Conference
at Aalborg University)
ISSN 1603-9696
www.trafikdage.dk/artikelarkiv

Vehicle Delay-Driven Passenger Delay Modelling – An Agent-Based Copenhagen Case Study

Mads Paulsen, madsp@dtu.dk
PhD Student, DTU Management Engineering

Thomas Kjær Rasmussen, tkra@dtu.dk
Assistant Professor, DTU Management Engineering

Otto Anker Nielsen, oani@dtu.dk
Professor, DTU Management Engineering

Abstract

Travel time of passengers is often uncertain due to lack of punctuality of public transport services. Whereas Automatic Vehicle Location (AVL) data makes it easy to measure the punctuality of public transport vehicles themselves, calculating door-to-door passenger punctuality is challenging, as both the intended and realised routes of passengers have to be considered. This study introduces a MATSim mesoscopic simulation framework for evaluating passenger punctuality caused by vehicle delays in the railway system in the metropolitan area of Copenhagen. Based upon empirical train delay data for 65 weekdays in the autumn of 2014, the model shows that passenger punctuality is considerably smaller than train punctuality, with 17.8% of the passengers using the railway system being delayed more than a minute compared to their intended plan.

Introduction

In most public transport networks a substantial share of passengers use more than one line to satisfy their transport needs. When transferring between such lines there is a risk that the passenger may not reach the desired run of the next line in time due to delays of the previous line. Such delays play a vital role when dealing with passenger delays [1].

Paradoxically, public transport operators are generally evaluated on the basis of their vehicle punctuality and reliability, not the door-to-door punctuality of passengers [2]. Although some research has touched upon passenger perspectives in railway timetabling [3], incorporating such passenger reliability measures in the actual evaluation of public transport systems is not done in practice. The first step to address this is to be able to calculate passenger consequences when public transport systems fail to meet full punctuality.

Ongoing research seems to focus on inferring such passenger delays by the use of data from individual passengers. Passenger delays were calculated in [4] and [5] for a closed metro system in Shanghai using Automatic Fare Collection (AFC) data, and [6] combined smartphone data with AVL data to infer observed and intended trips through San Francisco's Muni network. Such studies can provide important information on passenger delays, but the methods require data that typically cannot be assumed to be available across the entire network in general public transport systems.

Instead, in this study, we present an agent-based model framework for calculation of passenger delays by means of vehicle delays retrieved from AVL data generally collected in such systems. Passenger delays are evaluated by comparing the door-to-door travel times of intended routes (according to the planned timetable) to realised travel times that are modelled on the basis of empirical train delays and a semi-adaptive route choice model, allowing passengers to intelligently choose which train to board on-the-go. A detailed dataset consisting of train delays from 65 weekdays is used as input for the model, allowing the analysis of disaggregate day-to-day variations in passenger delays as well as overall aggregate analyses. Such a model to infer passenger delays from vehicle delays should facilitate improvements of public transport systems towards better user experiences for the passengers using them.

Before the model and the case study is introduced, a short literature review of existing approaches to modelling passenger delays modelling is presented.

Background

Determining passenger delays is a complex task as it requires knowledge about both the intended and realised route for every passenger in the system. In timetable-based public transport networks the intended routes and their associated travel times can be modelled in numerous ways [7]. For instance by utilising a diachronic graph [8] a dual graph [9], [10] or a mixed line database [11]–[13]. However, routes found with the above methods are only valid when assuming full punctuality of services.

Unfortunately, full punctuality is rare, why dynamic adaptive route choice models are needed when modelling how passengers move through an actual unreliable schedule-based public transport network. [11]introduced (and models on a toy network) four adaptive strategies ranging from sticking to the same series of stop as intended (run/line choice (principle 1)) to choosing an optimal path based on full a priori knowledge of the entire (delayed) network (principle 4).

In the literature several of such adaptive route choice models for public transport have been proposed, both using frequency-based approaches (e.g. [14], [15]), and schedule-based approaches (e.g. [16]). Only a few of these deal with passenger delays explicitly.

[17] and [18] both found passenger delays by averaging over several simulated days with delays drawn from statistical distributions. However, in both studies full a priori knowledge of present and future delays was assumed (principle 4, [11]).

[19] presented a model with internally modelled queueing delays caused by passengers, but without including other causes of vehicle delays.

[20] used an approach with realised train delays and adaptive passenger route choices in the time dimenstion (run choice) similar to principle 1 of [11], but allowing full optimal adaptive route choice based on full information after a passenger is delayed above a certain threshold. The model was applied to the suburban railway network of Copenhagen using micro-simulated railway running times. [21]extended this by using realised delays provided by the railway operator.

[22] introduced a model for a part of the Dutch train network used for evaluating passenger effects of adding additional trains under disruption, but explicitly suggested using the model for analysing passenger delays in future work.

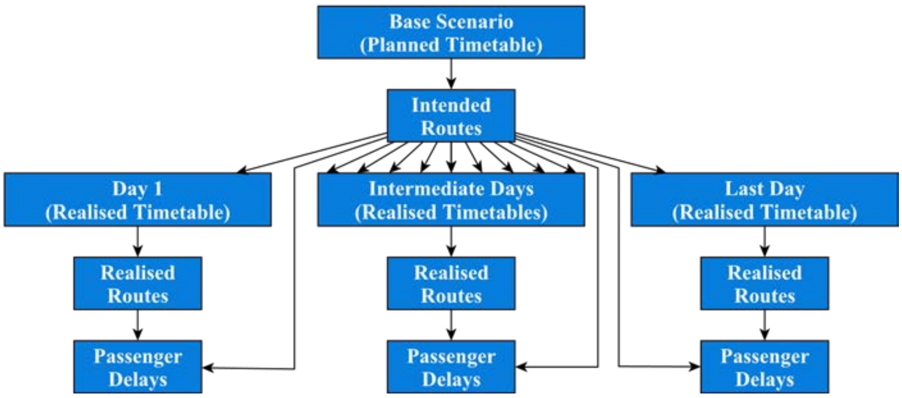

Figure 1 – The model framework consisting of a base scenario with a planned timetable providing intended routes, and scenarios for each historical day based on realised timetables resulting in realised routes.

Methodology

This study contributes to the literature by evaluating door-to-door passenger delays in a multi-modal public transport system of a large-scale metropolitan area using readily available train delay data. A diverse mix of transfers between low and high frequency lines for both buses and trains is secured through the multi-modality of the model, currently only represented in the literature when applying smartphone data of individual passengers.

The model firstly identifies planned routes according to the timetable for all public transport trips, and then imposes the realised timetable under an assumption that passengers make adaptive on-the-go choices in the time dimension (run choice) between stops/stations.

Such a process is repeated for each historical day with train delays available, allowing door-to-door travel patterns of each individual passenger to be modelled for each of these days. Comparing the trip travel times from these days to the base scenario where all trains run according the schedule allows the calculation of passenger delays. The approach is illustrated in Figure 1 and formally written in Algorithm 1 in the appendix.

The model is based on version 0.9.0 of MATSim [23] an open-source mesoscopic transport simulator capable of modelling door-to-door transport on an individual level for both public transport and car traffic. The events-based public transport router-extension of MATSim [24] allows public transport users to reach optimal planned routes through a day-to-day learning process by comparing the scores of performed routes across different iterations.

When the day is simulated, only adaptive choices in the time dimension (run choice) can be performed, in the sense that agents choose the first departing vehicle that can take them to the next stop in their planned route. This corresponds to principle 1 of [11].

Case Study and Results

The case study is based on a recently developed MATSim scenario for the metropolitan area of Copenhagen [25]. This includes the entirety of the public transport system spread across different varieties of trains (24 lines), buses (400 lines), metro (2 lines) and ferries (1 line), but also includes car users assigned to a high-resolution network of 379,907 unidirectional links.

A base scenario using the planned timetable as well as 65 historical days from the autumn of 2014 with realised timetables have been run with the model. The realised timetables were provided by the railway operator (DSB) and covers all train runs in the region (excluding metro).

A 1 % sample of the population with a total of 15,976 agents was used in the tests presented in this paper. 8,274 public transport trips were simulated per historical day, out of which 4,086 trips used the railway system at some point during the trip.

Table 1 – Proportion of passengers and trins arriving early, on time, and late.

	Passengers	Trains
Delay ≤ -1 minute	5.7 %	11.6 %
|Delay| < 1 minute	76.6 %	77.0 %
Delay ≥ 1 minute	17.8 %	11.4%
Average Delay [min]	1.49	0.15
SD of Delay [min]	8.10	2.33
Number of Observations	265,362	2,362,880

The aggregate results shown in Table 1 illustrate that the average delay for trips using the railway system is 1.49 minutes and that the standard deviation of the delay is 8.10 minutes. It is seen that the majority of passengers (76.6 %) arrives within one minute of their intended arrival time. However, a considerable number of passengers (17.8 %) are delayed more than a minute, whereas only a few passengers arrive more than a minute earlier than expected (5.7 %).

The proportion of passengers arriving late is in line with findings in [21] that showed that 15 % of passengers arrive late when only considering the suburban railway network of Copenhagen. Both the average delay and proportion of passengers arriving early were much larger in that study though (7.5 - 8.4 minutes and 17.3 - 19.6 %, respectively).

As expected, it is also seen the train punctuality is considerably higher than the passenger punctuality. This is even more evident in Figure 2 where the empirical cumulative distribution functions of the passenger and train delays of the current study are illustrated. About 5 % of all trips are delayed more than 10 minutes, whereas arriving much earlier than intended only occurs rarely.

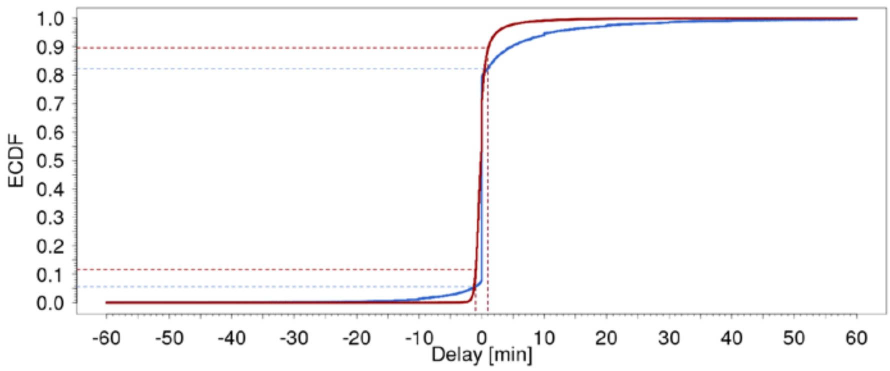

Figure 2 – Empirical cumulative distribution function of delaus and train arrival delays at each station.

Finally, Figure 3 shows the mean and standard deviation of delays based on the location of origins and destinations, respectively. The most evident tendency seems to be that trips departing from the city center are more volatile than trips heading for the city center. This seems plausible as passengers travelling from the city center most likely will use a high frequent (and likely irregular) service before transferring onto a service with a lower frequency, which they may miss due to the irregularity of the first service.

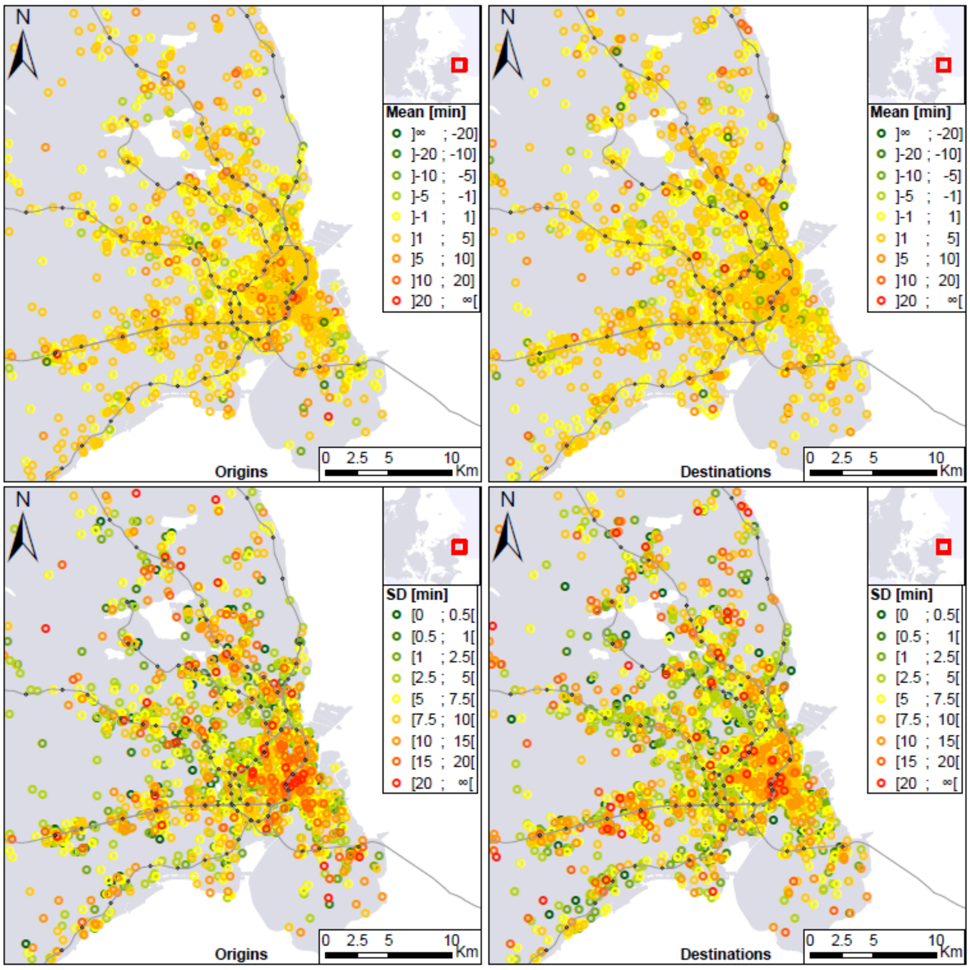

Figure 3 – Mean (top) and standard deviations (bottom) of trips going from (left) to (right) specific locations.

Conclusions

This study has presented an agent-based framework for determination of passenger delays in large-scale multi-modal transport systems based on train delays from AVL data. In the application to the metropolitan area of Copenhagen, the model required no additional empirical passenger information, and was still able to calculate door-to-door delays of individual agents across different modes of the public transport system. Results from these initial tests showed that both the mean and variance of passenger delays were notably higher than for vehicle delays, and that the distribution of passenger delays had an overweight of positive delays.

Future work includes increasing the population sample to provide a better coverage of the model area. This would allow carrying out detailed analyses such as investigating the effect of transfer types (to/from bus/train) and stations. Additionally, the study can be extended to include calculation of optimal route choices for each historical day, i.e. route choice under the assumption of full knowledge of all present and future delays. This is currently not possible for the realised timetables due to memory limitations, but by allocating more resources (than the 128gb currently available), this extension should indeed be possible. As the current adaptive route choice only considers the time dimension, this extension could provide an (admittedly optimistic) estimate of the lower bound of the passenger delays.

References

[1] J. Bates, J. Polak, P. Jones, and A. Cook, "The valuation of reliability for personal travel," *Transp. Res. Part E Logist. Transp. Rev.*, vol. 37, no. 2–3, pp. 191–229, Apr. 2001.

[2] R. B. Noland and J. W. Polak, "Travel time variability: A review of theoretical and empirical issues," *Transp. Rev.*, vol. 22, no. 1, pp. 39–54, Jan. 2002.

[3] J. Parbo, O. A. Nielsen, and C. G. Prato, "Passenger Perspectives in Railway Timetabling: A Literature Review," *Transp. Rev.*, vol. 36, no. 4, pp. 500–526, Jul. 2016.

[4] Z.-B. Jiang, F. Li, R.-H. Xu, and P. Gao, "A simulation model for estimating train and passenger delays in large-scale rail transit networks," *J. Cent. South Univ.*, vol. 19, no. 12, pp. 3603–3613, Dec. 2012.

[5] Y. Sun, J. Shi, and P. M. Schonfeld, "Identifying passenger flow characteristics and evaluating travel time reliability by visualizing AFC data: a case study of Shanghai Metro," *Public Transp.*, vol. 8, no. 3, pp. 341–363, Dec. 2016.

[6] A. Carrel, P. S. C. Lau, R. G. Mishalani, R. Sengupta, and J. L. Walker, "Quantifying transit travel experiences from the users' perspective with high-resolution smartphone and vehicle location data: Methodologies, validation, and example analyses," *Transp. Res. Part C Emerg. Technol.*, vol. 58, pp. 224–239, Sep. 2015.

[7] Y. Liu, J. Bunker, and L. Ferreira, "Transit Users' Route-Choice Modelling in Transit Assignment: A Review," *Transp. Rev.*, vol. 30, no. 6, pp. 753–769, Nov. 2010.

[8] A. Nuzzolo and F. Russo, "An equilibrium assignment model for intercity transit networks," in *Proceedings of TRISTAN II Conference*, 1994.

[9] J. Añez, T. De La Barra, and B. Pérez, "Dual graph representation of transport networks," *Transp. Res. Part B Methodol.*, vol. 30, no. 3, pp. 209–216, Jun. 1996.

[10] O. A. Nielsen and R. D. Frederiksen, "Large-scale schedule-based transit assignment - further optimization of the solution algorithms," *Oper. Res. Comput. Sci. Interfaces Ser.*, vol. 46, pp. 119–144, 2009.

[11] M. D. Hickman and D. H. Bernstein, "Transit Service and Path Choice Models in Stochastic and Time-Dependent Networks," *Transp. Sci.*, vol. 31, no. 2, pp. 129–146, May 1997.

[12] C. O. Tong and S. C. Wong, "A stochastic transit assignment model using a dynamic schedule-based network," *Transp. Res. Part B Methodol.*, vol. 33, no. 2, pp. 107–121, Apr. 1999.

[13] O. A. Nielsen, "A stochastic transit assignment model considering differences in passengers utility functions," *Transp. Res. Part B*, vol. 34, pp. 377–402, 2000.

[14] G. Gentile, S. Nguyen, and S. Pallottino, "Route Choice on Transit Networks with Online Information at Stops," *Transp. Sci.*, vol. 39, no. 3, pp. 289–297, Aug. 2005.

[15] F. Teklu, "A Stochastic Process Approach for Frequency-based Transit Assignment with Strict Capacity Constraints," *Networks Spat. Econ.*, vol. 8, no. 2–3, pp. 225–240, Sep. 2008.

[16] A. Nuzzolo, U. Crisalli, and L. Rosati, "A schedule-based assignment model with explicit capacity constraints for congested transit networks," *Transp. Res. Part C Emerg. Technol.*, vol. 20, no. 1, pp. 16–33, Feb. 2012.

[17] O. A. Nielsen, C. H. Overgård, and A. Daly, "A Large-scale model system for the Copenhagen-Ringsted railway project," in *Travel Behaviour Research*, 1st ed., D. Hensher, Ed. Oxford, UK: Elsevier, 2001, pp. 603–626.

[18] A. Nuzzolo, F. Russo, and U. Crisalli, "A Doubly Dynamic Schedule-based Assignment Model for Transit Networks," *Transp. Sci.*, vol. 35, no. 3, pp. 268–285, Aug. 2001.

[19] M. H. Poon, S. C. Wong, and C. O. Tong, "A dynamic schedule-based model for congested transit networks," *Transp. Res. Part B Methodol.*, vol. 38, no. 4, pp. 343–368, May 2004.

[20] A. Landex and O. A. Nielsen, "Simulation of disturbances and modelling of expected train passenger delays," in *Computers in Railways X*, 2006, vol. 1, pp. 521–529.

[21] O. A. Nielsen, O. Landex, and R. D. Frederiksen, "Passenger delay models for rail networks," in *Schedule-Based Modeling of Transportation Networks*, Boston, MA: Springer US, 2009, pp. 27–49.

[22] Y. Zhu and R. M. P. Goverde, "Dynamic passenger assignment during disruptions in railway systems," in *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 2017, pp. 146–151.

[23] M. Balmer, K. Meister, M. Rieser, K. Nagel, and K. W. Axhausen, "Agent-based simulation of travel demand: Structure and computational performance of MATSim-T," *2nd TRB Conf. Innov. Travel Model. Portl. Juni 2008*, no. June, pp. 1–33, 2008.

[24] S. A. Ordóñez, "New Dynamic Events-Based Public Transport Router," in *The Multi-Agent Transport Simulation MATSim*, London: Ubiquity Press, 2016, pp. 126–136.

[25] M. Paulsen, "Bus and Passenger Simulation within a Combined Agent-Based Multi-Modal Assignment Model," Technical University of Denmark, 2016.

Appendix

Algorithm 1

1: Create the planned timetable, S^P.
2: **while** not converged **do**
3: Run MATSim with the events-based public transport router-extension ([24]) using S^P to get the intended travel times, t_T^I, for all trips, $T \in \mathcal{T}$.
4: Create realised timetables, T_D^R, for all historical weekdays, $D \in \mathcal{D}$.
5: **for** all days, $D \in \mathcal{D}$ **do**
6: Simulate D in MATSim using the realised timetable, T_D^R, and the intended routes from 3, while allowing agents to make adaptive run choices in order to obtain the realised travel times,t_T^R, for all trips, $T \in \mathcal{T}$.
7: **for** all trips, $T \in \mathcal{T}$ **do**
8: Find the passenger delay of T, d_T, as the difference between the intended travel time from 3 and the realised travel time from 7,

$$d_T = t_T^R - t_T^I.$$