Denne artikel 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

Data-driven estimation of walking and waiting times for transfers in multi-modal public transport

Morten Eltved[1], Niklas Christoffer Petersen[1,3,*], Philip Lemaitre[2,4]

[1] Technical University of Denmark

[2] IT University of Copenhagen

[3] Trafikselskabet Movia

[4] Rejsekort & Rejseplan

Abstract

Transfers and connections between lines in a public transport network are a major part of the planning of good and reliable timetables. But it is difficult to observe whether the planned connections are used by the passengers. This research focuses on transfers from busses to trains and utilises smart card and automatic vehicle location data to estimate the walking and waiting times at these transfers. By appliying Bayesian inference for estimation of walking time distributions it is furthermore possible to quantify the uncertainty of a transfers planned by the traffic planners are possible in the real world.

1 Introduction

Transfers are unavoidable in many public transport journeys and the passengers consider this part of the journey as very unattractive (Schakenbos et al. 2016). The attractiveness of the transfer is highly dependent on the walking and waiting times and whether the passenger misses a connection, since this can result in highly increased journey times (Dixit et al. 2019). This study focus on estimating the walking time distributions for transfers, allowing timetable planners to plan more effective and thereby attractive correspondences between services. To study the transfer time distributions we combine automatic fare collection (AFC) data from the Danish smart card (Rejsekort) and automatic vehicle location (AVL) data from busses and trains.

Earlier work on transfer time distributions by Wahaballa et al. (2018) and Wahaballa et al. (2019) had the advantage of knowing whether the passenger performed a shopping transaction during a transfer, but since most AFC systems do not include this information, we investigate other methods for determining the walking time distribution with and without the passenger performing an activity during the transfer.

2 Model

To model the behaviour of transferring with or without an activity during the transfer, we propose to use a hierarchical mixture model. A transfer without activity is a transfer done with normal walking speed from the alighting to the boarding stop. It is assumed that all transfers without activity from alighting to boarding stop uses the same direct path, making transfer time independent of the path and only dependent on walking speed of the passenger and distance from alighting to the boarding stop. The transfer without activity can then be modelled by a skew-normal distribution describing variance in walking speed and distance effect on the direct walking time from alighting to boarding stop.

The transfer with activity is a transfer, where an activity affects the transfer time, such as running to catch the train, buying coffee on the way or stopping to tie their shoes. The transfer with activity is modelled by a uniform distribution to accommodate the different types of activities, which can either increase or decrease the transfer time, given the following model:

$$P(\alpha, \sigma, \lambda | T) \approx \prod_{k=1}^{K} \Big[\prod_{i=1}^{N} (\lambda_k \mathsf{SN}(T_{i,k} | \mu_k, \sigma_k, \alpha_k) + (1 - \lambda_k) U(T_{i,k} | 0, 1)) \Big]$$

There are K different pairs of bus to train platforms, N observed transfers times, T is the observed transfer time and λ describes the share of transfer without activity to transfer with activity. The parameters μ and σ is the mean and standard deviation of transfer without activity walking speed, where α control the skewness of the distribution.

3 Case study

A large transfer station in the Copenhagen area, Valby Station, is used for testing the model. An overview of the station is shown in Figure 1. In the travel planner for public transport in Denmark, Rejseplanen, the bus stops are considered as one stop and all train platforms are considered one stop and the walking time between the stops is scheduled to 4 minutes. For the initial tests only passengers transferring from the bus stops A and B to platform 3 are considered.

3.1 Data sources and pre-processing

Data from September 2019 is used for the analysis, and our experiment combines three different data sources:

- Bus AVL data: arrivals and departures of busses to/from bus stop points.
- Train AVL data: arrivals and departures of trains to/from train platforms.
- AFC data: tap-ins and tap-outs, either performed inside a bus, or on a train platform.

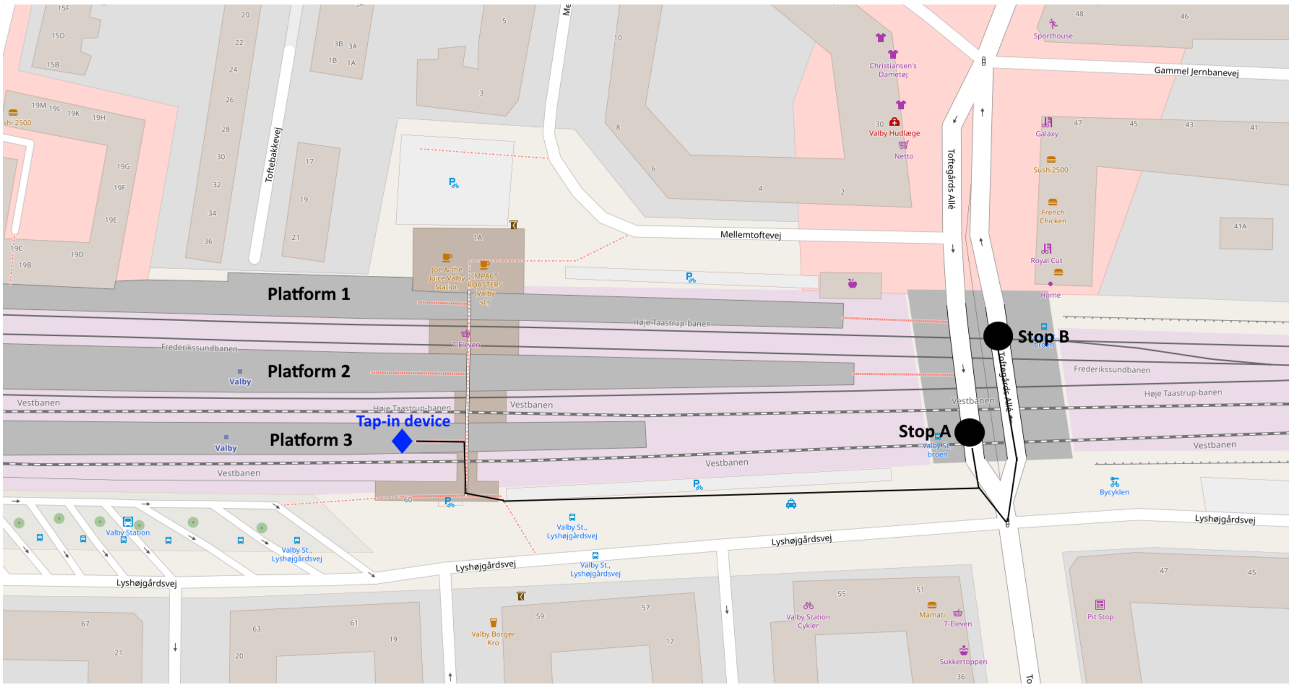


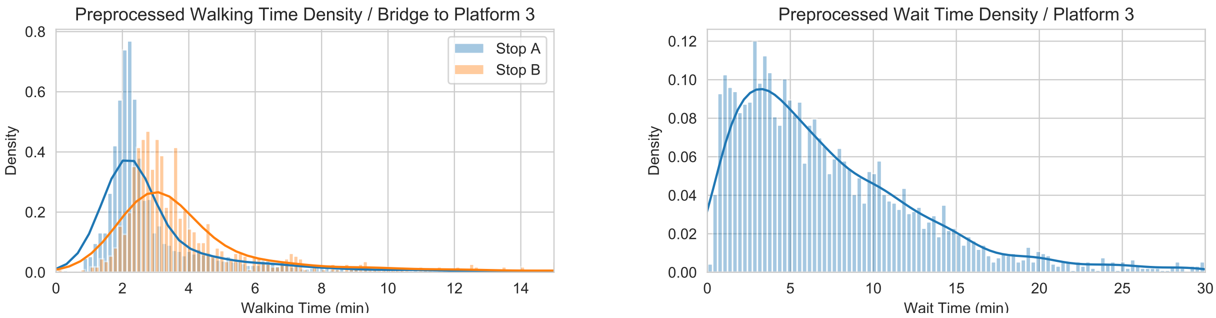
Figure 1: Overview of experiment location - Valby Station

In this AFC-system, each passenger is required to tap-in at all transfers, and to tap-out when the trip is finished. Since the AFC equipment is installed physically on the busses, the AFC and Bus AVL data can be directly related without any issues. On the other hand the train AFC equipment is installed on the train platforms, and not in each individual train.

Matching AFC data with Train AVL data can thus introduce uncertainty, as a passenger might have had several choices of trains given the observed tap-in/out time pattern. To minimise this uncertainty we only consider trips that was finalised on the next train station, and only for 3 selected destination stations. We match data based on the simple rule: tap-in at the transfer site should be before the train departed the transfer station, and tap-out at the destination should be after the train arrived at the destination station. If multiple trains fits the rule, we choose the one that arrived nearest the tap-out time. However, with the basic rule more than 70% of passengers are matched to a single feasible train, since the average headway of the services is around 8 minutes. This passenger-to-train assignment will be further developed to possibly include probabilistic measures such as the one developed in Zhu et al. (2017). Figure 2 shows the distributions of the two main variables in the pre-processed and combined dataset: Walking Time and Waiting Time. The final dataset consists of $2\,995$ matched observations.

3.2 Results of case study

The model is estimated by Bayesian inference using NUTS sampling with 3000 iterations and 4 chains. The results are shown in Table 1. The share of transfers without activity is estimated to 80% from stop A to platform 3 and 83% from stop B to platform 3 (hereafter stop A and B). All transfers without activity from stop A is possible within the scheduled walking time of 4 minutes, compared to the transfers without activity from stop B where 17% are above the scheduled walking time. The mean transfer time without activity is 1 minute higher for stop B than stop A with a lower confidence interval difference of 30 seconds and upper confidence interval difference of 1.6 minutes. The fitted predictive posterior of transfers without activities is shown in Figure 3



(a) Walking time distribution

(b) Waiting time distribution

Figure 2: Walking time to/Waiting time at train platform 3

(Left), and including transfers both with and without activities (Right).

Parameters	mean	sd	hpd 2.5 %	hpd 97.5%	Above scheduled walking time	ess	$\hat{R}$
μ_A	1.86	0.02	1.83	1.89	-	4304.0	1.0
μ_B	2.58	0.05	2.49	2.68	-	2931.0	1.0
σ_A	0.72	0.02	0.67	0.76	-	3826.0	1.0
σ_B	1.08	0.05	0.99	1.19	-	3088.0	1.0
λ_A	0.80	0.01	0.78	0.82	-	3918.0	1.0
λ_B	0.83	0.02	0.79	0.86	-	4133.0	1.0
α_A	0.98	0.02	0.95	1.00	-	3833.0	1.0
α_B	0.95	0.06	0.84	1.00	-	2164.0	1.0
Transfer Without Activity A	2.25	0.60	1.08	3.45	0 %	3724.0	1.0
Transfer Without Activity B	3.16	0.91	1.48	5.06	17 %	3999.0	1.0
Predictive Transfer A	4.87	6.59	0.00	22.64	17 %	4185.0	1.0
Predictive Transfer B	5.13	5.64	0.05	20.53	30 %	4058.0	1.0

Table 1: Posterior means and statistics in minutes.

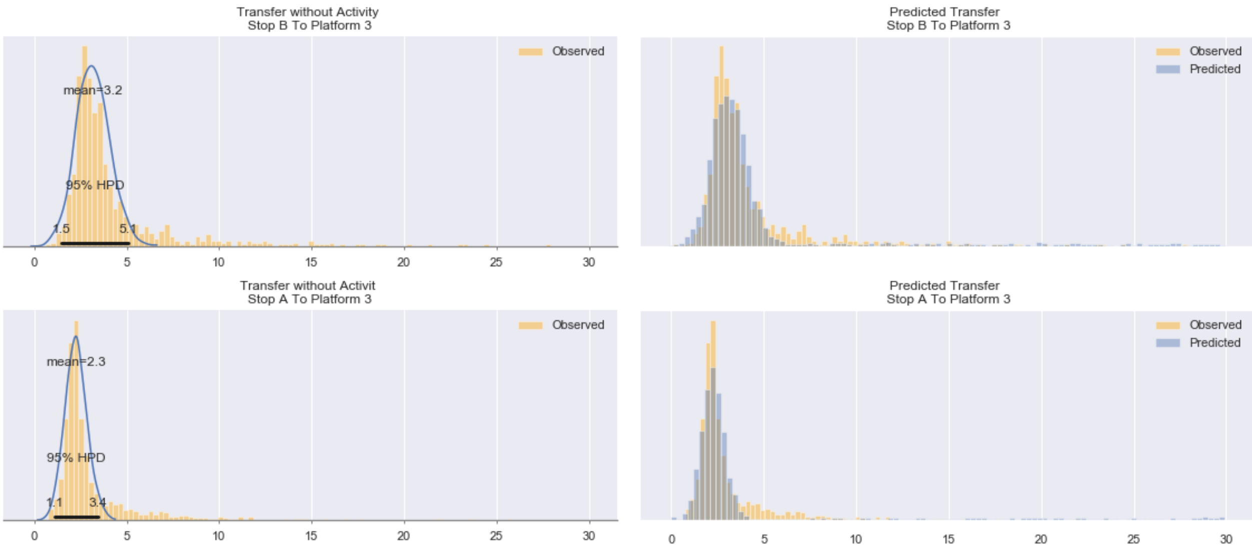


Figure 3: Predictive posterior of transfer without activity and transfer in minutes.

4 Conclusion

This study presents a method and model to determine the true walking time and waiting time distribution by adjusting the walking time gained from combining raw AFC and AVL data. The proposed hierarchical mixture model explicitly distinguish between transfers with and without activities, in order to extract the true walking time distribution. By applying the method on the case study, we have shown that 17 % are not able to make the specific transfer from Stop B to Platform 3 within the scheduled walking time.

The results show how our proposed method can be used to 1) identify problematic over-optimistic scheduled transfer times, and 2) identify connections, where a reduction in the scheduled walking times is possible, and this can be used to further optimise the effectiveness of the public transport network.

Since the method does not only yield point estimations for walking and waiting times, but complete distributions, the method can also quantify the probability for a specific passenger will make a given connection. This can e.g. be used in real-time and user centered ITS applications, allowing the user to specify and customize the uncertainty threshold for trips with transfers.

References

Dixit, M., Brands, T., van Oort, N., Cats, O. & Hoogendoorn, S. (2019), 'Passenger Travel Time Reliability for Multimodal Public Transport Journeys', Transportation Research Record 2673(2), 149–160.

Schakenbos, R., La, L., Nijenstein, S. & Geurs, K. T. (2016), 'Valuation of a transfer in a multimodal public transport trip', Transport Policy 46, 72–81.

Wahaballa, A. M., Kurauchi, F., Schmöcker, J.-d., Hemdan, S. & Iwamoto, T. (2019), Transit Travel Time Distributions Estimation Based on Passive AFC Data, in 'Transit Data 2019', Paris, France.

Wahaballa, A. M., Kurauchi, F., Schmöcker, J.-D. & Iwamoto, T. (2018), Rail-to-Bus and Bus-to-Rail Transfer Time Distributions Estimation Based on Passive Data, in 'CASPT 2018', Brisbane, Australia.

Zhu, Y., Koutsopoulos, H. N. & Wilson, N. H. (2017), 'A probabilistic Passenger-to-Train Assignment Model based on automated data', Transportation Research Part B 104, 1–21.