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 transfer walking time distribution in multimodal public transport systems based on smart card data

Morten Eltved[1,5], 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
[5] MOE Tetraplan

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 can be reached with the planned walking time between stops. This research focuses on transfers from busses to trains and utilises data from Rejsekort and automatic vehicle location data to estimate the walking times at these transfers. By applying Bayesian inference for estimation of walking time distributions it is furthermore possible to quantify the uncertainty of a transfers planned by the traffic planners. The method is applicable to all bus to train transfers and is a convenient way to obtain robust walking times, which take into account extra time used for walking due to stairs and waiting at traffic signals.

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

Figure 1 illustrates a transfer site, and the overall terminology for the proposed method. The goal is to estimate the walking time distributions for the different path pairs (4 shown), without explicit knowledge of passengers true walking time nor knowledge on whether or not they performed an activity during their transfer.

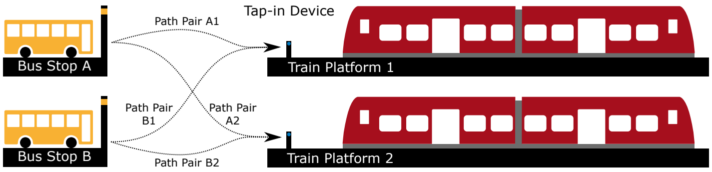


Figure 1: Overview of challenge and infrastructure setup.

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 of the bus to the train platform (refered to as a path pair). 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 distributions for direct walking time (e.g. without activity) can then be modelled by a beta-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 beta-distribution to accommodate the different types of activities, which can either increase or decrease the transfer time, given the following model:

$$P(\boldsymbol{\Omega}, \boldsymbol{\lambda} | \boldsymbol{W}^O) \propto \prod_{q=1}^{Q} \left[\prod_{i=1}^{N_q} \left[\lambda_q P(W_{q,i}^O | \Omega_q^D) + (1 - \lambda_q) P(W_{q,i}^O | \Omega_q^A) \right] \right] P(\boldsymbol{\Omega}^D) P(\boldsymbol{\Omega}^A) P(\boldsymbol{\lambda}).$$

There are Q different pairs of paths from bus to train platforms, each with N_q observed transfers times, $W_{q,i}^O$ is the i'th observed transfer time for path pair q and λ_q describes the share of transfer without activity to transfer with activity for the path pair. The vectors $\boldsymbol{\Omega}^D$, $\boldsymbol{\Omega}^A$ are parameters for the beta distributions and controls the mean and shape for transfers without activity, respectively transfer with activity. The model is illustrated as an probabilistic graphical model in Figure 2.

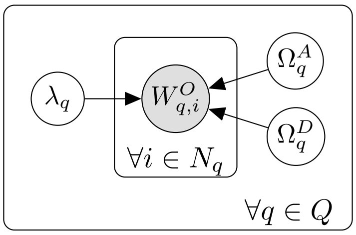


Figure 2: The model represented as a probabilistic graphical model.

3 Case study

Our case study is conducted for the entire Eastern Denmark for November 2019. We include most train stations serviced by DSB, metro stations and some local train stations. Figure 3 shows a map of the included stations. The model was estimated on 129 stations with a total of 1,009 path pairs. Only path pairs with 100 or more observations during November 2019 were estimated, as these pairs then have an average of at least three transferring passengers pr. day. The final dataset consists of 542,713 observations, i.e. unique transfers.

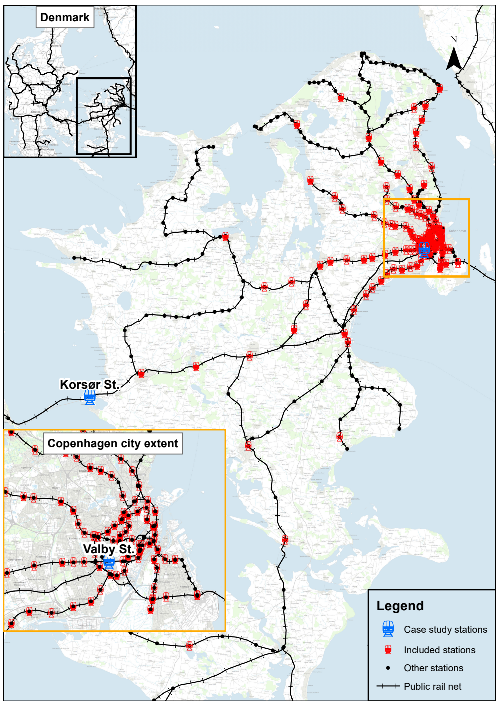
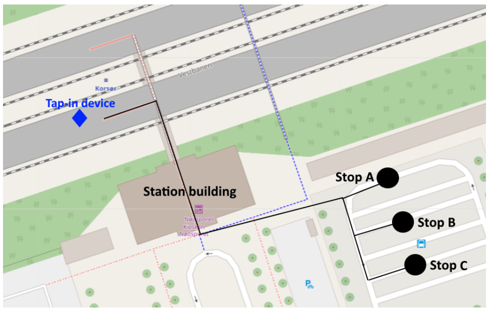


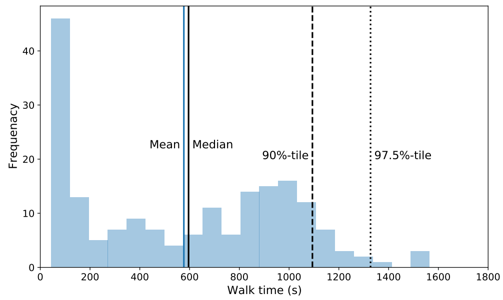
Figure 3: Overview of included stations in the analysis. Background map source: GeoDanmark (2020)

3.1 Results from specific station (Korsør)

Korsør Station has several different bus stops, which are located close to the station entrance as shown in Figure 4a. In Figure 4b the observed times from when the bus arrives at Stop B until the passenger taps in at the platform is shown. This shows that a majority of passengers take around 800-1200 seconds to walk, which seems unreasonable as there are also several passengers who are able to walk to the platform in less than 200 seconds. This might be due to bad coordination between bus and train, which allows the passenger to wait in the station building before going to the platform. These passengers thus have an activity during the transfer, which must be accounted for when estimating the time needed to walk from bus stop to train platform.



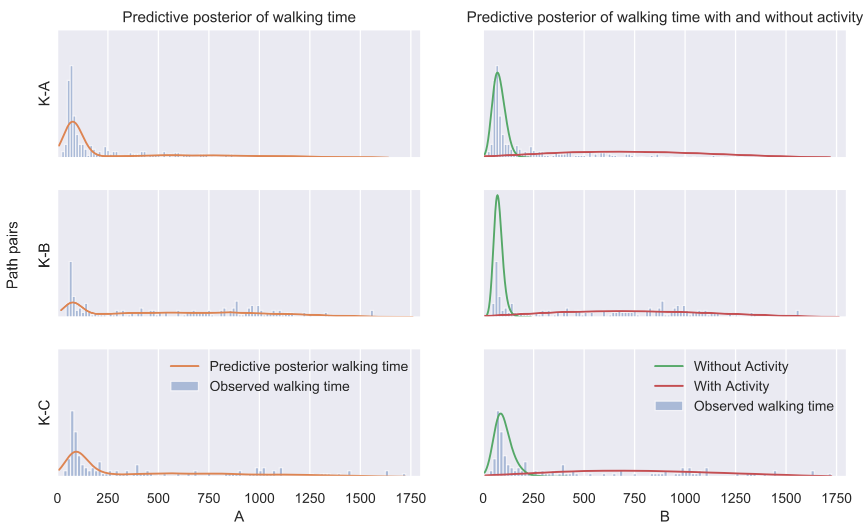
(a) Overview of Korsør Station



(b) Observed walking times (from alighting the bus until tap-in at platform) from AFC data at Korsør Station (rural station on Zealand, Denmark), specifically path pair connecting bus stop B to the platform (K-B) which has an approximate network walking distance of 150 meters.

Figure 4: Overview of station layouts with background map from OpenStreetMap and observed walking times from Rejsekort data

The estimated walking time for the three path pairs from bus stop to train platform is shown in Figure 5. In the left column the resulting joint distribution for walking times is shown. This shows, that the necessary walking times would be overestimated if the raw data from Rejsekort is used directly to fit the distribution for walking time. In the right column the resulting distributions from using the proposed model is shown. The green lines showing distributions for passengers assumed to walk directly seem to follow more correctly the passengers who seem to walk directly to the train platform in less than 200 seconds. The distribution for the passengers assumed to have an activity during the transfer covers all the passengers who must be assumed to have done an activity during the transfer.

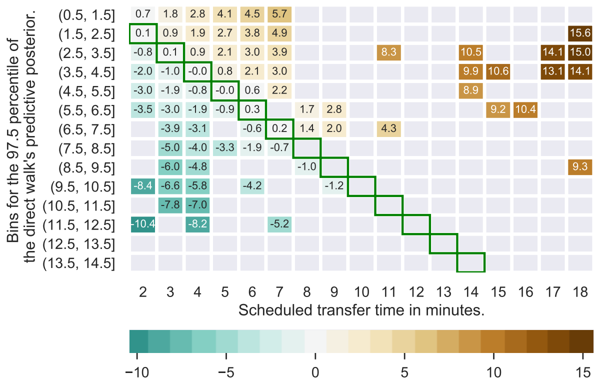


(A) The predictive posterior walking time distribution is generated from the weighting of the direct walking share of activity and direct walking time. (B) The distribution of activity and direct walking time distribution without the direct walking share.

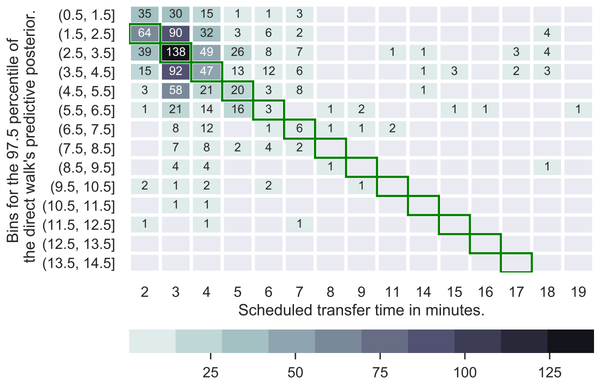
Figure 5: Korsør station - Predictive posterior of walking time compared to observed walking time.

3.2 Results for all stations

The overall results of the analysis for all 129 stations is visually illustrated in Figure 6, showing the average difference between scheduled time used in Rejseplanen and the 97.5^{th} percentile of the estimated transfer time distribution (6a), respectively the number of path pairs (6b). The majority of the path pairs' scheduled time are between two and five minutes (6b), which is also reflected in a the large number of path pairs being estimated in this interval. However, there are also several path pair with a scheduled walking time above 14 minutes which are estimated to take 2-5 minutes and also some path pairs estimated to take above 8.5 minutes with a scheduled walking time at 2-5 minutes. The path pairs with a scheduled walking time above 14 minutes are all scheduled with a too high walking time compared with the estimated walking time with an average difference ranging from 8.9 to 15.6 minutes. The majority of path pairs with an estimated walking time above 8.5 minutes have a too low scheduled walking time. Their estimated walking time have an average difference below -4.2 minutes of the scheduled walking time. The figures in Figure 6 makes it easy for the transit agencies to identify path pairs, which needs further investigation to ensure that most passengers can catch their train when transferring from bus to train.



(a) Average difference between scheduled and estimated walking time.



(b) The count of the number of path pairs for each comparison of estimated and scheduled walking time.

Figure 6: Comparison of scheduled and estimated walking of the path pairs of the 129 stations. The green squares indicated where the scheduled walking time is within 30 seconds of the estimated walking time's 97.5^{th} percentile.

4 Conclusion

This study has presented a novel methodology for providing accurate walking time distributions at transfers from bus to train based on smart card data. The model requires AVL data from busses and smart card data where the passenger must tap-in at the train station, preferably at the platform to avoid uncertainty of possible time spent in a station building.

The proposed approach is able to reproduce the observed times from the passenger alights a bus until tap-in at the platform using a hierarchical Bayesian mixture model, where passengers are assumed to either walk directly to the platform or perform an activity during the transfer. The model is applied to a large-scale case study with 129 stations in the Eastern part of Denmark.

The model can be easily applied at scale, and thus offer a more feasible methodology to obtain walking times at transfers than manual surveys where passengers are followed through the transfer and also provide more detailed information than using a point estimate for the walking time based on walking distance. The distribution for walking time for the direct walking passengers can be compared to the scheduled walking time published by public transport agencies, and thereby identifying places where extra scheduled walking time is needed. In this way the agencies are able to plan more reliable connections between busses and trains.

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.

GeoDanmark (2020), 'Background map from GeoDanmark'.
URL: https://www.geodanmark.dk/

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.