Monitoring the performance of the pedestrian transfer function of train stations using automatic fare collection data

J.P.A. van den Heuvel \textsuperscript{a,b,*}, J.H. Hoogenraad \textsuperscript{c}

\textsuperscript{a}NS Stations, Stationshal 17, 3511 CE Utrecht, The Netherlands
\textsuperscript{b}Department of Transport \& Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands
\textsuperscript{c}Spoorgloren, Schapenhoeve 11, 3992 PL Houten, The Netherlands

Abstract

Over the last years all train stations in The Netherlands have been equipped with automatic fare collection gates and/or validators. All public transport passengers use a smart card to pay their fare. In this paper we present a monitor for the performance of the pedestrian function of train stations which is based on data from the automatic fare collection system. To our knowledge this is the first study that uses smart card data in the context of pedestrian behaviour at train stations. To illustrate the added value of the monitor, various applications for a number of train stations are presented.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).
Peer-review under responsibility of Department of Transport \& Planning Faculty of Civil Engineering and Geosciences Delft University of Technology

Keywords: pedestrian; train station; automatic fare collection; smart card; data

1. Introduction

Insight in passenger flows is essential information for the development and operations of train stations. Without this information, it is very difficult to set capacities of pedestrian facilities of the terminal (ie. stairs and escalators) correctly, and to position station services (ie. travel information, ticketing and shops) at the right position, as perceived by train passengers. From its Station Experience Monitor Netherlands Railways (or "NS") has learned that passenger experience tends to be better when the pedestrian function of a station is performing better. Negative factors are for example congestion at platform exits (escalators and stairs), difficulties in orientation or lack of travel information at the right place.

Until recently, pedestrian flows at Dutch train stations have been estimated using data for boarding, alighting and transferring passengers for an annual average workday, enriched by data from questionnaires which are held throughout the year. In case of reasonable doubt in the estimated outcomes or in case of a need in more accurate and/or detailed information, manual counts at specific stations were organized. This has proven to be a costly method of data

\textsuperscript{*} Corresponding author. Tel.: +31-6-29598119.
\textit{E-mail address:} j.p.a.vandenheuvel@tudelft.nl

2352-1465 © 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).
Peer-review under responsibility of Department of Transport \& Planning Faculty of Civil Engineering and Geosciences Delft University of Technology
doi:10.1016/j.trpro.2014.09.107
collection. Moreover, with data from occasional manual counts it is hardly possible to infer trends in passenger flows, causes of traffic peaks and the time spent at the station.

Recently, NS’ division NS Stations has been developing three new tools to generate more and more detailed information about passenger flows at train stations. The first is the Station Transfer Model, which has been developed in 2010. This model translates traffic data already available at NS’ train operating division into estimates of pedestrian flows at platforms, escalators and stairs (Van den Heuvel et al. (2012)). Although the model is based on data for an annual average workday and a planned train schedule, the temporal and spatial allocation of train passengers has resulted in a significant improvement in the usability for station design activities. The second tool is SMART Station, which has been in development since 2011. It combines Bluetooth, WIFI and Infrared technology to measure pedestrian flows inside the terminal (Van den Heuvel et al. (2014); Ton (2014)). This measurement tool has proven to deliver sufficiently accurate and detailed data on an hourly basis for both station design and operations. Upfront costs of deployment are relatively high, because SMART Station implies the installation and configuration of a large number of sensors the semi-outdoor space of a train station. When installed marginal costs of measurements are very low. Therefore, SMART Station is a cost-effective tool for large, complex stations with continuous operational challenges or redevelopment activities.

The third tool taps into the automatic fare collection (AFC) data which is increasingly available at NS’ train operating division. In the 2012 Aurora project, NS has developed an algorithm which is designed to predict the number of passengers inside specific trains using AFC-data (Hoogenraad et al. (2013)). By adding the spatial characteristics of train stations to the existing AuroraROCKT-tool, we have created a tool which is capable of estimating pedestrian flows and dwell times at all train stations where NS is the train operating company. This AuroraROCKT-Station tool currently delivers data for the Station Transfer Monitor which is presented in this paper. The next section introduces the concept of station transfer performance. Section 3 describes the tool, which is an extension of the existing Aurora algorithm. Several applications of the tool are presented in section 4. This paper ends in section 5 with conclusions.

2. Defining station performance

Similar to its other policies, guidelines and monitors, NS has linked the performance of the pedestrian function of train stations - referred to as “station transfer” - to the concept of The Pyramid of Customer Needs. In his PhD-thesis, Van Hagen (2011) has adapted Maslow’s hierarchy of needs to train stations and passengers as its main users.
Time is a key element at the bottom of this pyramid, which refers to it as "speed", as Van Hagen (2011) states: "the majority of customers choose as short a travel time between origin and destination as possible". The door-to-door travel time can be defined by the travel time to and from the station by the access and egress modes, the in-train time and time spent at the train station. Excluding time spent at station activities (ie. meeting, getting a coffee), the time at the train station is the total of walking and waiting time. The Station Transfer Monitor is focussed at this last part of total travel time: the walking and waiting time at the train station, before, after and during a train trip, or in short "the station dwell time".

For the Station Transfer Monitor we have categorized station dwell times of train passengers in departures, arrivals, transfers and activities. Observed behavior at train stations has shown that departing passengers tend to spend more time at the station than arriving passengers (Van Hagen (2011); Van den Heuvel et al. (2014); Ton (2014)). Departing passengers built in buffer time before a train trip to avoid the risk of missing the preferred train service. For others the arrival time at the station is constrained by the schedule of the access mode. This results in a wide distribution of dwell times of boarding passengers of each train. Arriving passengers tend to exit the station as quickly as possible, either to get to their scheduled egress mode on-time or to get to their destination as quickly as possible. This results in a narrow distribution of station dwell times of alighting passengers of each train. The dwell time of passengers who change trains at the transfer station depends on the difference between arrival and departure times of both trains. Although increasing, a small share of the station visitors uses the station itself as origin or destination for their activity. Our monitor covers the three largest groups: departures, arrivals and transfers.

3. Extracting information from smart card data

In the past fifteen years, smart card-based fare collection systems have replaced traditional fare collection systems in many public transport worldwide (Fleishman (2003)). This trend has been accompanied by the exploration of the potential of smart card systems as source of data. Many public transport operators started to search for opportunities to improve operations (Furth et al. (2003); Bagchi and White (2005)). A still increasing number of researchers started to develop algorithms to be used for extraction of transit times (Hofmann and O’Mahony (2005)), Origin-Destination matrices (Zhao (2004); Cui (2006); Zhao et al. (2007); Wang (2010); Chan (2007)), train occupation data (Kusakabe et al. (2010); Hoogenraad et al. (2013)), travel times and route choices in networks (Sun and Xu (2012); Ma et al. (2013); Sun and Schonfeld (2014)), transit service reliability data (Cham (2006)) and transfer patterns (Jang (2010)). In their recent state-of-the-art article Pelletier et al. (2011) have classified the the use of smart card data in public transport, using three classes: strategic long-term network planning, customer behaviour analyses and demand forcasting, tactical analyses of schedule adjustments, longitudinal and individual patterns, and operational supply-demand indicators and AFC-systems operations. We have found no work that links AFC-data to pedestrian behaviour at stations.

In order to understand the application of smart card data for the measurement of station performance, some design choices of the AFC system in The Netherlands need to be explained. NS accepts smart cards (OV-Chipkaart in Dutch) as a payment means since 2010. The AFC system is based on a distance-based fare, as the smart card can be used to travel in the entire country. Passengers are required to both check-in and check-out in order to determine the correct fare. In the train system, the check-in and check-out processes are organized at the train stations, to keep the boarding and alighting process of the trains as efficiently as possible. At all stations, gates and/or validators are placed at the entrances, deliberately as far away from the platforms as possible. This allows minimization of investment and operating costs of AFC-equipment, it minimizes the risk of crowding due to a better spatial distribution of pedestrian flows, and creates a comfortable, secured travel domain inside the station. Transferring passengers are not required to check-out and check-in at their transfer station as long as they transfer between trains of the same operator. At the end of 2013 about 30% of the passengers use a smart card to pay their fare. This fraction will increase quickly, since NS has ended the acceptance of paper train tickets on 9 July 2014.

While implementing the AFC system, NS has been developing the AuroraROCKT algorithm to estimate train occupations using AFC data. A key objective of these algorithms is to correctly combine the realized time table (train times and actual platforms) with check-in and checkout-pairs from the smart card data. Data from the in-train inspection teams were used to validate and optimize this process.

Extending the AuroraROCKT-algorithm with information about the locations of gates and validators allowed distribution of the check-in/out-pairs of each train at each platform in time and space. The AFC-data generates check-in/out
times and gate/validator ID’s at both departure and arrival station. The realized time table data delivers the exact time and platform of all trains which potentially could be used, given the time and station of each check-in/out-pair. The in-train inspection team data deliver the information to distribute check-in/out pairs to the correct train, in case multiple train services for an check-in/out-pair are viable travel options. Finally, the data on gate/validator locations delivers the data on which station entrances have been used for each check-in/out-pair. This sequence delivers for each check-in/out pair the departure station, including entrance used, and departure time and platform, and similar data for the arrival station. On an aggregate level this data allows us to estimate station dwell time and the distribution of passengers over the various platforms and station entrances.

4. Cases

4.1. Measuring the efficiency of the pedestrian function at Utrecht Central station

The first example aims at assessing the performance of the pedestrian transfer function in terms of speed. As shown by the Pyramid of Customer Needs, this factor is a key element of station experience. Because arriving passengers arrive by train in large groups and tend to leave the station as quickly as possible, monitoring the check-out times of arrivals is a good proxy for transfer speed.

In our example we use data from Utrecht Central station (Fig. 2). This is the largest station in The Netherlands in terms of number of train passengers: 170,000 departures and arrivals, and 60,000 transfers per average workday. The station consists of fourteen platform tracks, of which some are used by through trains, and some are used by terminating trains. The stopping position of through trains tends to be centered right under the station hall, close to multiple platform exits. Due to historical track capacity increases, terminating trains tend to stop at one of the platform sections, concentrating passenger flows at just one platform exit. Particularly platform 18/19 is used in that way. The A-section (18A and 19A) is near the station hall, while trains at the B-section stop over 100 meters south of the station hall. Passengers from these trains need to walk between 100 and 300 meters before they arrive at the platform exit.

Fig. 3 shows the cumulative percentage of passengers that have checked out within a given time frame after arrival of their train at a specific platform. In the left graph we have presented the cumulative distribution functions for the platforms which are marked in Fig. 2. In the right graph we have presented a distribution over time of the day for a selection of these platforms. The underlying dataset consists of more than 2,000 smart card transactions per trace of arrivals during off-peak hours on weekdays in December 2012. On the horizontal axis, the time is shown between train arrival and the time of check-out. The cumulative number of passengers that has checked-out is shown at the vertical axis. Each trace in the graph represents a distribution function for one platform.

As expected for the left graph, the distribution functions for platform 18B and 19B stand out. For the other platforms with centralized stopping positions of the trains, it takes about a minute for the fastest passengers to check-
out at the smart card validators which are located at the station exit. In contrast, even the fastest passengers of platform 18B/19B reach the station exit only after 3 minutes. After about 5 minutes after train arrivals, 90% of the passengers from most platforms have left the station. Again, the slowest passengers from platforms 18B and 19B take about two minutes more to leave the station.

For peak hours, we expect longer exit times due to additional delays when crowding at platform exits occurs. Fig. 3b confirms this expectation for platform 19B where it takes about an additional minute for the slowest passengers to check out during peak hours. For platform 7 arrivals additional delays due to crowding do not occur because of an optimal train stopping position with respect to the platform exits.

The data shows that a small fraction of the passengers check-out later than the 10 minutes. This effect is mostly related to passengers who might have returned to the station for a next train trip, after having left the station after the previous one without having checked out. This issue will decrease as soon as passengers are more getting used to the smart card, and when the gates at the gated stations have been closed. Although to much lesser extend, this effect is also related to passengers who stay at the station for some time after their train trip, for example to buy something to eat or to meet.

The data also shows that a very small fraction (< 1%) of the passengers has checked-out before the train has arrived. This is related to an algorithm issue. For some routes, passengers can chose from a direct train service which takes a little longer, and a faster trip with a transfer. In this case, the algorithm incorrectly assigns the passengers to a direct train that arrives a few minutes later, resulting in a negative check-out time after train arrival. Currently, we are looking into ways to improve the algorithm to resolve this issue.

Currently NS scans all stations annually using the Station Tranfer Monitor to detect the platforms with long check-out times. The results have been used for further investigation. Some findings already have resulted in changes in train operations, station operations, or even in the station layout.

4.2. Assessing the waiting time effects of train schedule changes for transfers at Deventer station

The second example illustrates the use of the Station Transfer Monitor to assess the impact of train schedule changes on waiting times for passengers who change trains at a specific station (transfers). For this case, the example of Deventer station is used (Fig. 4). In terms of number of train passengers Deventer has the 29th largest station in The Netherlands: 20,000 departures and arrivals, and 8,000 transfers per average workday. The station is situated at a crossing of a north-south main line (Zwolle-Roosendaal) and a east-west main line (Amsterdam/Schiphol-Enschede). Both intercity and local trains run half-hourly off-peak on the east-west line, increasing to a 15-minute interval to the west during peak hours. The north-south corridor has a all-day 30-minute interval intercity service.

In December 2012, the train schedule in the Netherlands has significantly been changed after the opening of a new rail link between Lelystad and Zwolle (Hanzelijn). This resulted in a parallel east-west main line from Amsterdam-
dam/Schiphol to Zwolle. For large numbers of passengers travelling between the west and north of The Netherlands, this line improved travel times. However, on a few relations the transfer times increased due to shifts in the arrival and departure times of trains at the east-west corridor. Due to the schedule structure change, it turned out not to be possible to maintain the optimal train connection times between both main lines at Deventer station. This resulted in an increase of the waiting time for transferring passengers.

Fig. 5 shows the transfer times for passengers using a smart card for several weekdays at non-peak hours. In the 2012 schedule, 50% of the transferring passengers had a connection within 8 minutes, and about 90% within 15 minutes. After the train schedule change (from 10 December 2012) 50% of the transferring passengers had a connection within 15 minutes, and a significant fraction more than 20 minutes.

In order to alleviate the waiting time at Deventer, NS has improved the station services at the platform. Except by a small Kiosk at the central side, the central platform building had not been in use since the last tenant had left in 2010. Currently, the building hosts a combined coffee bar and small book-store, which offers a convenient waiting facility for transferring passengers.
Amsterdam Central station is used for the third example. As the second-largest station in the Netherlands, it is a busy hub with 170,000 departures and arrivals, and 15,000 transfers per average workday. The station has three large entrances at the city center side, and two station entrances at the waterfront side. These five entrances are connected to the platforms by three transfer tunnels: the West tunnel (W), Central tunnel (C) and East tunnel (E). The sixth entrance is located at the west side of the station and goes directly to platform 1/2A. Most trains stop either at the west side of a platform (A-section) or on the east side (B Section). A very limited number of trains stop centrally, in a way the central platform section is used. When a train stops at the A-section the majority of the train doors are located west of the West tunnel. When a train stops at the B-section, the train doors are located between the Central tunnel and East tunnel, and further to the east. When a train stops centrally, the platform section between the East tunnel and West tunnel is used. Choices on the train position are made to maximize track capacity, because all tracks at this station are being operated at capacity during peak hours (Starmans et al. (2014)).

In this example, we focus on the route choice of arriving and departing passengers between the three station entrances at the city center side. We have analysed passengers’ choices for check-in (departures) or check-out (arrivals) locations inside the station building using smart card data from December 2012 (Fig. 7). The tables in the lower part of the figure show the distribution of passengers over the three station exits at the city center side for all trains that stopped at either the A-sections or the platforms, or the B-sections.

Fig. 7 shows that the West tunnel is the preferred route for both arriving and departing passengers for A-section trains. Only 28% departing passengers and 16% of the arriving passengers use one of the other transfer tunnels. We can conclude that at Amsterdam Central the majority of the passengers for A-section trains uses the transfer tunnel which is closest to the platform section from which their train runs. B-section trains show a different picture. The majority of the departing passengers still prefers the West tunnel, despite the distance from the location where the trains run. For the arriving passengers both the Central and East tunnel are mostly preferred, although the West tunnel still has a significant share here as well. Further data analyses for platforms, time of the day (ie. peak and off-peak hours), day of the week (ie. weekday and weekend) did not reveal any differences in the pattern described above.

Various factors could be at work here. The first is the transfer tunnel quality. The West tunnel is a wide tunnel with station services and retail, which impacts station experience in a positive way. The East tunnel is narrow, outdated tunnel, without retail, and with limited station services due to lack of space. The central tunnel has been under construction due to the construction of a new metro line underneath. Both attributes impact station experience negatively. For a more discussion on the impact of station experience and route choice we refer to Verhoeoff (2014). The second factor is the location of access and egress modes outside the station. Although the various modes of transport are spread around the station, the majority is located at the west.
Our analysis of passenger route choice in Amsterdam Central station illustrate that planners should be careful with assumptions about pedestrian route choice inside train stations. The common assumption that the majority of arriving passengers takes the first transfer tunnel on their way to the station exit, is confirmed from the smart card data. The assumption that arriving and departing passengers equally are distributed over the various routes through the station does not hold. Moreover, there still is a significant difference between the A-sections and the B-sections. The share of B-section arrivals who uses the West tunnel is 3.5 times as large as the share of A-section arrivals who uses the East tunnel. So the common assumption of symmetrical pedestrian flows does not hold. These findings have had significant implications for a major overhaul that is being planned (see Starmans et al. (2014) for more details).

5. Conclusions

We have developed a Station Transfer Monitor based on automatic fare collection data to monitor the performance of the pedestrian function of train stations in The Netherlands. As fas as we are aware of, this is the first time that AFC-data is applied to pedestrian flows and dwell times at train stations.

The Station Transfer Monitor offers NS a very detailed and easily deployable monitoring tool to measure and monitor important parameters of pedestrian functions of train stations. The three cases presented in this paper have given insight in the potential value of smart card data for operations and design of train station. The case of Utrecht Central station has shown that smart card data can help to detect transfer bottlenecks which cause inconvenience for arriving passengers. The case of Deventer station has shown the potential impact of train schedule changes on transfer times, and the importance of station services to alleviate waiting time. The case of Amsterdam Central station has given insight in commonalities and differences in chosen routes inside the train station by arriving and departing passengers, and the impact of train stopping locations, transfer quality and configuration of access and egress modes.

For this study we have used smart card data from a relatively small population out of the total train passenger population (about 30%). This could raise the issue whether our findings are applicable in other cases. With the rapidly increasing use of the smart card for all public transport in The Netherlands, our smart card data will soon represent the majority of train passengers. Repeating the analyses on more recent data will remove the risk of a sampling bias. Moreover, when the gates at the gated stations have been closed, the quality of the station dwell time of arrived passengers will improve. Finally, further improvement of our algorithms will decrease the impact passengers being...
assigned to the wrong trains in situations where multiple train services could have been used, based on the check-in and check-out pairs.

Regarding future research, we are looking deeper into the station dwell times of departing passengers. We have indications that these are at least partially related to the access mode choice. Bus passengers tend to be more constrained by the arrival times as set by the bus schedule, while pedestrians and cyclists have more degrees of freedom in determining their arrival time at the station before train departure. With the currently available smart card data it is not possible to analyse inter-modal transfers, so this research direction requires us to be able to combine data from different modes of public transport which in The Netherlands are often operated by different operators.

References

Jang, W., 2010. Travel time and transfer analysis using transit smart card data. Transportation Research Record - Journal of the Transportation Research Board 2144, 142–149.