Schedule-based estimation of pedestrian demand within a railway station

Flurin Hänseler

Nicholas Molyneaux

Michel Bierlaire

Amanda Stathopoulos

Transport and Mobility Laboratory, EPFL

May 2014

STRC | **14th Swiss Transport Research Conference**

Monte Verità / Ascona, May 14-16, 2014

Transport and Mobility Laboratory, EPFL

Schedule-based estimation of pedestrian demand within a railway station

Flurin Hänseler, Nicholas Molyneaux, Michel Bierlaire, Amanda Stathopoulos
{flurin.haenseler,nicholas.molyneaux,michel.bierlaire,amanda.stathopoulos}@epfl.ch

May 1, 2014

Abstract

A framework is outlined for estimating pedestrian demand within a railway station which takes advantage of the train timetable and train frequentation data, as well as various direct or indirect indicators of demand. These may include e.g. link flow counts, measurements of density and travel times, or historical information. The problem is considered in discrete time and at the aggregate level, i.e., for groups of pedestrians associated with the same origin-destination pair and with the same departure time interval. The formulation of the framework allows for a wide applicability to various types of railway stations and input data. A preliminary case study analysis of Lausanne railway station provides an example of such an application.

Keywords

Demand estimation, pedestrian flows, schedule-based estimation, public transport

1 Introduction

The demand for mobility is increasing at a fast pace, and so is the volume of traffic in general. Taking the network of the Swiss Federal Railways (SBB) as an example, the number of daily transported passengers has grown by approximately 50% in the last decade alone (Amacker, 2012). To cope with such a growth, SBB's transportation system has been continuously expanded over the past century. In particular, the frequency and the capacity of trains have been increased. However, one component that has largely been neglected in this context is that of rail access installations (Schneider, 2012). The capacity of pedestrian facilities simply has not been a limiting factor for a long time. Today, the latter are increasingly considered a bottleneck of the railway system, both in Switzerland and around the world.

During normal operation of a railway station, it is mostly large incoming and outgoing trains that lead to a high usage of pedestrian facilities. Following a train arrival, a potentially large number of passengers disembark, and then move as a dense crowd through the station. These 'pedestrian waves' may provoke congestion in platform access ways. Similarly, prior to train departures, outbound passengers typically cumulate on platforms which serve as waiting areas. If platforms are narrow, or the number of prospective passengers high, space quickly gets scarce. These two exemplary phenomena – pedestrian waves induced by train arrivals and the cumulation of outbound passengers on platforms – may have a negative impact on customer satisfaction, as well as the performance and safety of a train station.

By increasing the capacity of rail access installations, these negative effects can be alleviated. Unfortunately, the required investment is often very expensive. The example of Lausanne railway station, which is going to be extended in the next years, may be used to illustrate why: A planned enlargement of underground access ways, and a necessary widening of platforms require almost a complete reconstruction of the station. Furthermore, several buildings are under historical preservation protection and can hardly be altered. The biggest cost driver, however, is the requirement that the railway station be fully functional during the complete period of construction.

Given the complexity and cost of an expansion of rail access installations, a diligent planning and dimensioning is indispensable. Key in this process is the assessment of the usage of a railway station, i.e., the estimation of pedestrian demand. This demand is subject to significant temporal variability. Whenever a train arrives and pedestrians disembark, demand is high, and may be low otherwise. Such fluctuations are referred to as 'micro-peaking' (Hermant, 2012).

In this work, we aim at developing a methodology for estimating pedestrian demand in a railway station, taking into account the particular demand pattern induced by arriving and departing trains. An explicit integration of the train timetable allows to quantitatively appreciate its influence on pedestrian traffic in rail access installations. In the long run, this framework may be used to optimize the train timetable or track assignment in that context.

2 Literature Review

Pedestrian behavior in railway stations increasingly attracts the attention of academic research. Broadly, it can be distinguished between empirical studies aiming at characterizing behavior, and those concerned with its mathematical modeling.

In an early study, Daly *et al.* (1991) investigate the relationship between speed and flow and between flow and travel time in various pedestrian facilities of London's underground system. Lam and Cheung (2000) examine several metro stations as well as pedestrian areas in a shopping center in Hong Kong. Differentiating by trip purpose, flow capacities are evaluated and flow-travel time functions are calibrated. Compared to the results from London, users of Hong Kong's mass transit system are found to be better at dealing with high levels of congestion, which is attributed to their smaller physique and sociological differences.

In a related study, Lam *et al.* (1999) investigate the train dwelling time and the distribution of pedestrians on platforms in two stations of Hong Kong's Light Rail Transit system. A behavioral analysis reveals that people are significantly less willing to board a train if the latter is congested, and if the journey to be made is longer. Also focusing on train platforms, Zhang *et al.* (2008) quantitatively describe the process of alighting and boarding in metro stations in Beijing. A cellular automaton model is developed, calibrated on empirical data and complemented with a thorough literature review. Various empirically observed behavior patterns can be reproduced with high accuracy. Pettersson (2011) investigates the behavior of pedestrians on railway platforms from an architect's perspective. In particular, the effect of signposts, availability of seats and entrances on the distribution of pedestrians along the platform is investigated at the example of a Swedish and a Japanese case study. Concrete recommendations are made regarding how a more homogeneous distribution along a train platform can be attained.

To assess the design of a railway station, it is important to be able to predict the routes taken by pedestrians. Several studies have been dedicated to this endeavor. Again at the example of a metro station in Hong Kong, Cheung and Lam (1998) investigate the route choice between escalators and stairways leading to a train platform. A relationship between flow and travel time is first established. This characteristic relationship is then used in a choice model allowing to predict the percentage of escalator-users for ascending and descending directions as a function of prevailing traffic conditions. In a similar context involving two Dutch stations, Daamen *et al.*

(2005) have collected route choice data by following passengers through the facility from their origin to their destination. Likewise, a route choice model is estimated allowing to predict the influence of level changes in walking routes on passenger route choice behavior. It is concluded that the various ways of bridging level changes such as ramps, stairs or escalators have a different impact on the attractiveness of a route. Finally, Hoogendoorn and Bovy (2004) develop a comprehensive model for pedestrian route choice and activity scheduling. In their framework, every route and activity schedule is associated with a cost. It is assumed that pedestrians choose their route and activities such that the perceived utility is maximized. The methodology is applied to a case study of a major Dutch transportation hub.

With the exception of the last study, the mentioned work concentrates mostly on individual aspects of pedestrian behavior in railway stations such as flow-travel time relationships or way finding. In the following, various studies are briefly presented that consider an integral modeling of pedestrian behavior in a railway station.

Lee *et al.* (2001) present one of the first model-based studies of pedestrian flows in a railway station in the scientific literature. For a major station in Hong Kong's metro system, origin-destination demand and travel times are collected using a large number of human observers. From this data, flow-travel time relationships are derived, which are used in a relatively simple, network-based pedestrian flow model. A comparison between empirical data and model prediction indicates a good performance of the model.

A very comprehensive modeling framework for pedestrian flows in railway stations is due to Daamen (2004). A multitude of models for describing the processes of queueing, boarding, alighting, waiting, walking as well as route and activity choice are proposed, and jointly implemented. The framework is very rich in details, representing an agent-based hybrid queueing network/link flow-model operating in discrete space. Various case studies across the Netherlands are considered.

Kaakai *et al.* (2007) develop a similar model, albeit at the macroscopic level. They consider both discrete processes such as the arrival and departure of trains, as well as continuous processes such as the fill-up of railway platforms by pedestrians awaiting a train, or pedestrian flows in walking facilities. The model is represented as a Petri net and applied to a French case study involving a railway station with a single platform. Hanisch *et al.* (2003) and Tolujew and Alcalá (2004) seem to qualitatively follow a similar approach, but do not provide a mathematical specification of their model.

Again at the microscopic level, Xu *et al.* (2014) develop a model describing pedestrian behavior in a Chinese metro station. The framework is entirely based on a queueing network, i.e., all processes including entering the railway station, passing ticket gates, walking and boarding are represented by queues. This assumption seems valid for situations with very high demand and a high level of congestion. The framework is applied to estimate the maximum service rate of a metro station, as well as to determine the optimal inlet rate at the entrance at which this capacity is attained.

There are several more studies of pedestrian flows in railway stations that concentrate rather on a high level of accuracy for specific applications than on a methodological contribution. Most of them pursue an agent-based approach and describe various local challenges such as the placement of access gates in Lisbon (Hoogendoorn and Daamen, 2004), the re-design of access ways in the Swiss capital of Bern (Rindsfüser and Klügl, 2007), the evacuation of a metro station in Beijing (Jiang *et al.,* 2009), the modeling of waiting areas in German railway stations (Davidich *et al.,* 2013) or the design of a new station in South Africa (Hermant, 2012).

While most of the previously mentioned studies use sophisticated models for describing various aspects of pedestrian behavior such as walking, waiting or boarding, the methods used to estimate pedestrian demand are quite simplistic. Many do not even specify how these estimates are obtained. Other studies rely on flow counts that are converted to origin-destination demand values based on simple rules of thumb, such as assuming a uniform demand over time. Very few studies take the train timetable explicitly into account, but if so, then typically only for individual platforms.

Despite a considerable interest in pedestrian behavior models for railway stations, there thus seems to be a lack of dedicated methods for estimating pedestrian demand. Ideally, such a methodology should be able to reproduce the demand micro-peaks mentioned previously, i.e., it should be able to provide time-dependent demand estimates by explicitly taking the train timetable into account. To do so, different approaches seem conceivable. For instance, in the context of a university campus, Danalet *et al.* (2014) propose an activity choice model based on WiFi traces and individual class schedules. However, for most applications involving train stations, disaggregate data is still unavailable, and instead it is more efficient to estimate origin-destination (OD) demand at the aggregate level.

For problems concerning car traffic, such dynamic OD demand estimation methodologies are already well established. Inspired by the seminal work by Cascetta *et al.* (1993), a large number of statistical methods have been developed in the last two decades (Bera and Rao, 2011). By building on these achievements, this study aims at providing a dedicated estimation methodology for pedestrian OD demand in railway stations.

3 Estimation framework

To estimate time-dependent OD demand, ideally all relevant, available information sources should be taken into account. In particular, this may include various kinds of measurements of demand and the train timetable. For that purpose, in the following an assignment mapping is established that defines the temporal and spatial relationship between the OD volumes and these sources of information. This mapping will subsequently be used for solving the demand estimation problem.

3.1 Representation of space, time and pedestrians

Let the period of analysis be divided into a set of discrete time intervals $\mathcal{T}$, where each interval $\tau = [t_\tau^-, t_\tau^+]$, $\tau \in \mathcal{T}$, is of uniform length $\Delta t = t_\tau^+ - t_\tau^-$. With regard to space, let the network of pedestrian facilities be represented by a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N}$ represents the set of nodes $\nu \in \mathcal{N}$, and $\mathcal{L}$ the set of edges $\lambda \in \mathcal{L}$ connecting them. The ensemble of nodes through which pedestrians enter and leave the pedestrian facility network shall be referred to as the set of centroids and be denoted by $C \subset \mathcal{N}$.

Any two centroids may be connected by a route $\rho \in \mathcal{R}$, defined as a sequence of edges $\rho = (\lambda_1^\rho, \lambda_2^\rho, \dots)$. A subnetwork $\mathcal{G}_\alpha = (\mathcal{N}_\alpha, \mathcal{L}_\alpha)$, with $\mathcal{G}_\alpha \subset \mathcal{G}$ is referred to as an area α.

Each pedestrian is associated with a specific OD pair $\zeta = (\nu_o^\zeta, \nu_d^\zeta)$, where $\nu_o^\zeta, \nu_d^\zeta \in C$. The set of all OD pairs shall be denoted by $\mathcal{Z}$. Furthermore, for each OD pair ζ, the set of connecting routes shall be denoted by $\mathcal{R}_\zeta$.

Based on the above representation of time and space, the concept of demand can be defined. Let $d_{\zeta,\tau}$ represent the number of users leaving the origin ν_o^ζ during time interval τ towards destination ν_d^ζ. Furthermore, let the corresponding time-space expanded vector be denoted by $d = [d_{\zeta,\tau}]$.

3.2 Assignment mapping

The key concept of the estimation framework lies in the assumption of an assignment mapping $M_q(d)$ that defines the relationship between the time-dependent OD volumes d and direct or indirect observations of demand q such that

$$q = M_q(d)d. \tag{1}$$

In a congested network, M_q generally depends on the unknown OD demand. There may be various demand indicators such as flow and density measurements, walking speeds, travel times, or a priori information such as historical counts. These quantities may be available for the whole network, or only for parts of it. Let the ensemble of these indicators be represented by $\mathcal{Q}$, with $q \in \mathcal{Q}$, and let the corresponding set of mappings be denoted by $\mathcal{M}$, with $M_q \in \mathcal{M}$.

In practice, the 'true' assignment mapping $M_q(d)$ is rarely known, and needs to be approximated by an estimate $\hat{M}_q(d)$. Similarly, an indicator q is typically afflicted with a measurement error, i.e., only its estimate $\hat{q}$ is available. Introducing a random error ϵ_q, equation 1 can be expressed as

$$\hat{q} = \hat{M}_q(d)d + \epsilon_q, \tag{2}$$

where typically a zero mean for ϵ_q is assumed (Cascetta *et al.,* 1993). The map $\hat{M}(d)$ is usually obtained by means of a dynamic traffic assignment (DTA) model. In the context of pedestrian flows in a railway station, processes such as route choice, walking and waiting, boarding and alighting, and more specific activities such as buying a ticket, shopping or eating in principle need to be considered.

Of major importance in the assignment mapping of most demand indicators is a route choice model. Let the fraction of pedestrians associated with OD pair ζ choosing route $\rho \in \mathcal{R}_\zeta$ during time interval τ be denoted by $\delta_{\rho,\zeta}$. These choice fractions are usually not directly observable, and under congested conditions, they depend on the demand. One way of determining them is by means of a random utility model. According to Cheung and Lam (1998), pedestrian route choice is similar to that for conventional road traffic, i.e., users will choose their desired routes according to the shortest travel time, shortest travel distance, or a combination of both. An important difference to car traffic however exists for movements in the vertical direction, where pedestrians also consider the effort in level changes (Cheung and Lam*,* 1998*,* Daamen *et al.,* 2005). Without specifying an actual choice model, we assume in the following that the demand-dependent route choice matrix $\Delta(d) = [\delta_{\rho,\zeta}(d)]$ is known.

Equation 1 shall be illustrated by considering two examples of demand indicators. Let us first consider the flow $f_{\lambda,\tau}$ on a link λ during time interval τ. Generally, the flow on a link depends on the time-dependent OD demand, the route choice probabilities, and on the mapping of route flows to link flows. Let the latter be denoted by $a_{(\lambda,\tau),(\rho,\kappa)}(d)$, representing the probability that a user associated with route ρ and departure time interval κ reaches link λ during time interval τ. To estimate these probabilities, knowledge about time-dependent travel times on the network are needed. We again assume that such knowledge is provided by some suitable DTA model. The number of users entering link λ during time interval τ is then given by

$$f_{\lambda,\tau} = \sum_{\kappa\in\mathcal{T}} \sum_{\zeta\in\mathcal{Z}} \sum_{\rho\in\mathcal{R}_\zeta} a_{(\lambda,\tau),(\rho,\kappa)}(d)\delta_{\rho,\zeta}(d)d_{\zeta,\kappa}. \tag{3}$$

If we denote by $f = [f_{\lambda,\tau}]$ the time-space expanded link flow vector, and if $A(d) = [a_{(\lambda,\tau),(\rho,\kappa)}(d)]$ represents the corresponding assignment that maps route flows to link flows, equation 3 can be expressed in matrix notation as

$$f = A(d)\Delta(d)d. \tag{4}$$

Indicators different from flow might be of interest. In the context of a railway station, knowledge of the density on a railway platform or within a congested walkway could be useful. Edie (1963) provides a generalized definition of pedestrian density applicable to a discrete representation of space and time as used in this work. For example, we may be interested in the time-mean average of number of users in area α during time interval τ. In the following, this quantity is referred to as occupation and denoted by $n_{\alpha,\tau}$. Let $s_{(\alpha,\tau),(\rho,\kappa)}(d)\Delta t$ represent the amount of time spent in area α during time interval τ by a user that embarked on route ρ during time interval κ. The average occupation in area α during time interval τ can then be expressed as

$$n_{\alpha,\tau} = \sum_{\kappa\in\mathcal{T}} \sum_{\zeta\in\mathcal{Z}} \sum_{\rho\in\mathcal{R}_\zeta} s_{(\alpha,\tau),(\rho,\kappa)}(d)\delta_{\rho,\zeta}(d)d_{\zeta,\kappa}, \tag{5}$$

or in matrix notation with $n = [n_{\alpha,\tau}]$ and $S(d) = [s_{(\alpha,\tau),(\rho,\kappa)}(d)]$ as

$$n = S(d)\Delta(d)d. \tag{6}$$

3.3 Train-induced pedestrian behavior

The notion of edge flow and area occupation may be used to integrate the train timetable in the demand estimation process. The relationship between train arrivals and resulting exit flows, or between the accumulation of pedestrians on platforms and train departures have been documented and investigated by several researchers.

Daamen *et al.* (2008) study the flows occurring at train doors during boarding and alighting, reporting door capacities for various types of rolling stock. Buchmüller and Weidmann (2008) investigate the flows on platform access ways caused by alighting train passengers. Following the same approach, Molyneaux *et al.* (2014) discuss the concrete example of Lausanne railway station. Characteristic for such train-induced pedestrian arrival flows is the lagged onset of the flow after the arrival of the train, the saturation at a given capacity flow rate, and a subsequent decay (figure 1).

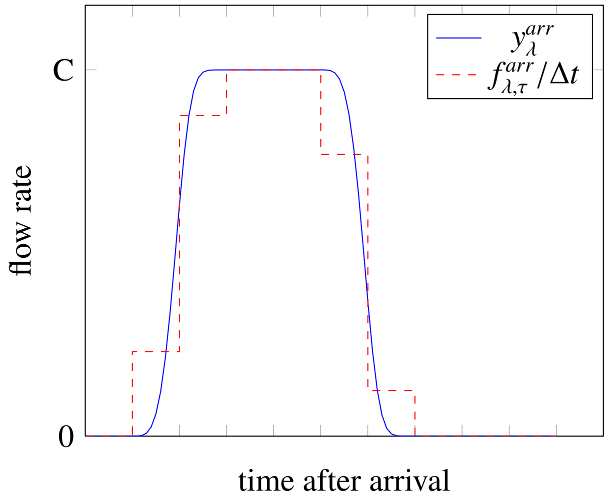

Figure 1: Illustration of the flow of alighting passengers in platform access ways after the arrival of a single train (adapted from Molyneaux *et al.,* 2014). Empirical evidence suggests that shortly after the onset of the flow, a capacity C is reached, which is maintained almost until all alighting passengers have left the platform. In dashed red, a discrete-time approximation $f_{\nu,\tau}$ is shown.

Departures of trains generate characteristic flow patterns as well. Tolujew and Alcalá (2004) and Hermant *et al.* (2010) study the waiting behavior of prospective train passengers on platforms. Based on qualitative observations, they assume that the accumulation of pedestrians on a train platform prior to the corresponding departure first follows an S-curve. Once the train has arrived, this accumulation is offset by pedestrians that start boarding, leading to a platform occupation as shown in figure 2.

Based on such empirical observations, a quantitative relationship between the train timetable and OD demand can be established. We assume that for each train m with arrival time t_m^{arr}, alighting volume Q_m^{al}, and a generic parameter vector $\gamma_{m,\lambda}^{arr}$, the arrival flow rate at (continuous) time t at link λ is given by

$$y_\lambda^{arr}(t; Q_m^{al}, t_m^{arr}, \gamma_{m,\lambda}^{arr}). \tag{7}$$

The specification of equation 7 may depend on the type of rolling stock, rail access installations, and other factors.

Similarly, we assume that the volume of passengers associated with a train m on the corresponding platform at time t depends on the boarding volume Q_m^{dep}, the departure time t_m^{dep}, as well as a

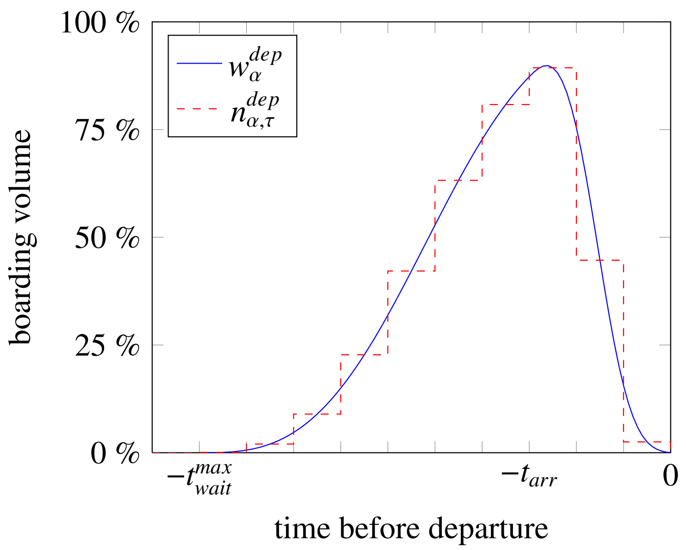

Figure 2: Illustration of the accumulation of outbound passengers on a platform prior to train departure (adapted from Tolujew and Alcalá, 2004). The earliest passengers arrive at a time t_{wait}^{max} before departure, and boarding starts shortly after the arrival of the train at t_{arr}, causing a drain of passengers.

generic parameter vector $\gamma_{m,\alpha}^{dep}$, given by

$$w_\alpha^{dep}(t; Q_m^{boar}, t_m^{dep}, \gamma_{m,\alpha}^{dep}). \tag{8}$$

During the time horizon $\mathcal{T}$, generally several trains arrive depart from each platform. Assuming a linear superposition, the cumulated train-induced flows of a total of M trains during time interval τ at link λ is given by (dashed red line in figure 1)

$$f_{\lambda,\tau}^{arr} = \int_{t=t_\tau^-}^{t_\tau^+} \sum_{m=1}^{M} y_\lambda^{arr}(t; Q_m^{al}, t_m^{arr}, \gamma_{m,\lambda}^{arr})\ \mathrm{d}t. \tag{9}$$

Accordingly, the occupation of outgoing passengers on a platform area α can be expressed as (dashed red line in figure 2)

$$n_{\alpha,\tau}^{dep} = \frac{1}{\Delta t} \int_{t=t_\tau^-}^{t_\tau^+} \sum_{m=1}^{M} w_\alpha^{dep}(t; Q_m^{boar}, t_m^{dep}, \gamma_{m,\alpha}^{dep})\ \mathrm{d}t. \tag{10}$$

3.4 Estimation of demand

The problem of estimating demand can be formulated as that of finding the vector d^* such that the various demand indicators q are matched at best by the corresponding mappings $M_q(d^*)d^*$.

If for each indicator q a suitable distance measure $\mathrm{dist}_q(\hat{M}_q(d)d, \hat{q})$ is defined, we have

$$d^*(Q, \mathcal{M}) = \underset{x \geq 0}{\arg\min} \sum_{q \in Q} \mathrm{dist}_q(\hat{M}_q(x)x, \hat{q}). \tag{11}$$

For an appropriate specification of the distance measures, the distribution of the error term ϵ_q in equation 2 should be taken into account. If for instance a normal distribution is assumed, the Euclidian norm can be used, such that equation 11 turns into a constrained, generalized least squares problem. Furthermore, depending on the correlation structure of q and d, the resulting problem may be equivalent to a maximum likelihood estimator (Cascetta *et al.,* 1993).

4 Case Study

To demonstrate the applicability of the presented demand estimation framework, and to discuss methodological and practical challenges associated with it, a case study of Lausanne railway station is currently in preparation. A description, yet without results, can be found in the following.

4.1 Description of site

Lausanne railway station is the largest railway station in French speaking Switzerland, serving over 120'000 train passengers every weekday (Amacker, 2012). Located at the junction of three national railway lines, it provides express train service to a variety of destinations across Switzerland and beyond. Passenger trains are primarily run by SBB, with additional international trains run by companies from neighboring countries. In total, there are over 650 trains arriving and departing from Lausanne every day. Across the train station square, the local metro system can be reached, and on the square itself several bus lines are accessible.

Figure 3 shows a schematic map of Lausanne railway station. The station encompasses the passing tracks #1–9 and the dead end track #70. Track #2 is used by freight trains and through traffic only, as it is not accessible by a railway platform. Platforms are named after the tracks they give access to, such as e.g. 'platform #3/4'. Except for platforms #1 and #70, all platforms are accessible from the city solely through two pedestrian underpasses, PU West and PU East. Furthermore, platform 9 is only accessibly from PU West. Longitudinally, the train station is divided into sectors A-D, where the historical ordering from East to West has been adopted.

According to an internal assessment by SBB, the peak demand over a workday in Lausanne is reached at around 07:45 when several long distance trains arrive and depart in close succession (Gendre and Zulauf, 2010). At this time of the day, more than 500 users might disembark during

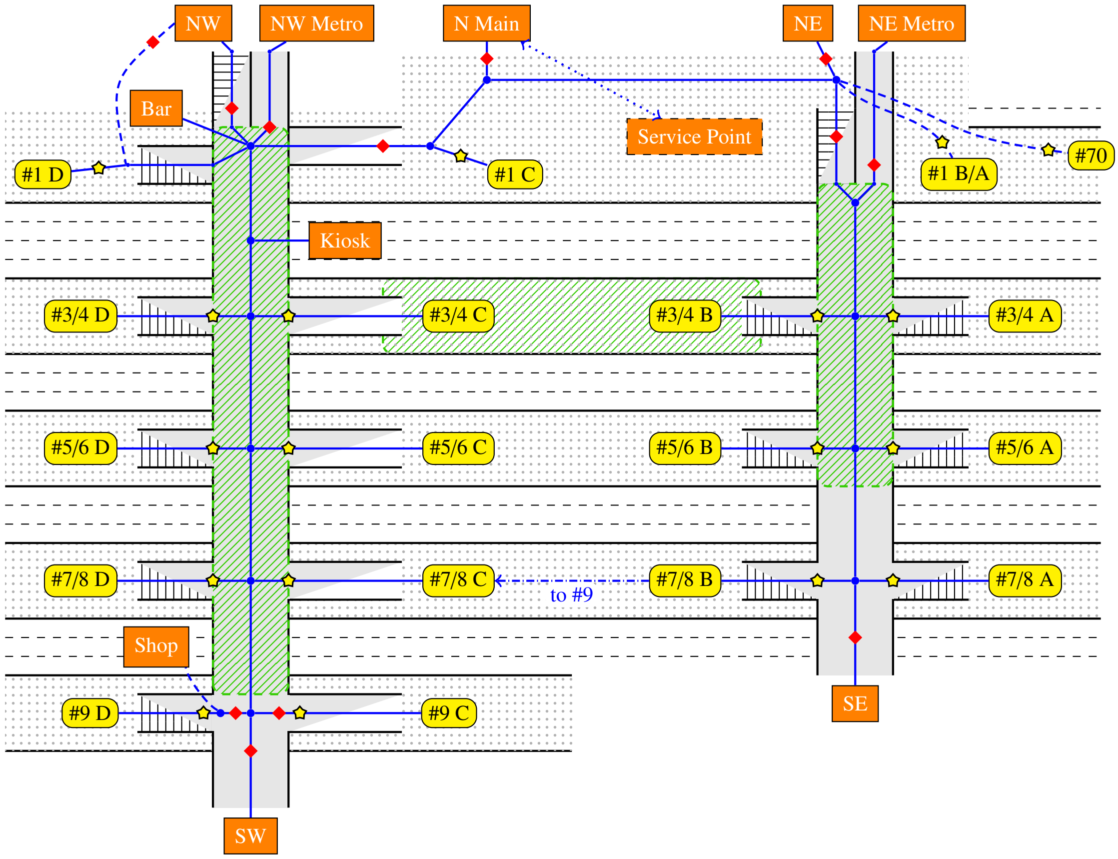

Figure 3: Schematic map of Lausanne railway station, providing ten tracks (#1–9, #70; denoted by dotted lines) that are served by platforms #1, #3/4, #5/6, #7/8, #9 and #70. Platforms are connected by two pedestrian underpasses referred to as PU West and PU East. Dotted areas denote joint walking/waiting areas, and grey areas represent walking facilities that are located underground. Under- and overground areas are connected by ramps and stairways (denoted by standard floor plan symbols). Each platform is represented by up to four centroids (yellow rectangles with rounded corners). Exit/entrance areas as well as three sales points are represented by further centroids (orange rectangles). The pedestrian walking network is represented by a blue graph connecting centroids and intersection nodes (solid blue circles). Dashed lines represent network links that cannot be directly shown on the scheme due to the chosen two-dimensional representation. Pedestrian counters are denoted by red diamonds. Facilities covered by a pedestrian tracking system are green shaded. Points at which train-induced passenger arrival flows are estimated are denoted by a yellow star.

a peak minute, whereas a few instants later it can be less than a hundred per minute (Alahi *et al.*, 2013). In fact, such a periodical concentration of pedestrians in stations is characteristic for the Swiss railway network that aims at bundling train arrivals and departures in order to minimize waiting time for transfer passengers in neuralgic railway stations (SBB-Infrastruktur, 2013). Besides fluctuations within a day, demand also varies from one day to another. Demand on weekends is typically lower than on weekdays. But even between similar days, such as two consecutive Mondays, demand can vary significantly (Hänseler *et al.*, 2013).

The pedestrian facilities in Lausanne railway station can roughly be divided into an East and a West side, each having a pedestrian underpass at its heart. In principle, it is possible to switch between the two sides by longitudinally crossing any of the main platforms. However, there is little need to do so, since all major platforms and exit ways are accessible from both PUs on a shorter way. The official signposting foresees only one cross-link on the North side for reaching platform #70 from PU West, and one on the South side through platform #7/8 for reaching platform #9 from PU East. In other words, the officially signposted network represents that of two parallel corridors that are connected at their extremities (figure 3).

Besides a distinction between East and West, pedestrian facilities can be classified with respect to the primary activity they provide for. In pedestrian underpasses and on platform access ways, pedestrians mainly walk. On train platforms, people mostly wait, board or disembark. In service areas such as restaurants or sales points, people eat or shop. Furthermore, there are some areas which cannot be unambiguously classified, such as e.g. the train station hall, where waiting, walking and other activities may happen concurrently.

In the morning peak hour, regular commuters with a good knowledge of the train station constitute the largest user group (Lavadinho, 2012). Given their familiarity with the location, they mostly walk along the shortest paths inside the station. Among all pedestrians, broadly four visitor types can be distinguished. This includes first and foremost inbound and outbound passengers, which arrive by train and leave for the city, and vice versa. Furthermore, there are local users, which take advantage of shopping facilities, or simply traverse the station, but never take a train. Last but not least, there are transfer passengers, which arrive and depart by train. During the morning peak hour, most visitors represent inbound passengers. Generally, the percentage of local users and transfer passengers is relatively low.

The classification into facility and visitor types might appear artificial, but turns out to be useful in the subsequent analysis. Indeed, in the present case study we consider pedestrian behavior in walking areas only. This choice is motivated by three reasons. First, walking areas represent the neuralgic part of the pedestrian facilities connecting the exterior of the railway station to the platforms. Second, in Lausanne only for these areas a data set is available for estimation and validation. Third, the modeling of pedestrian behavior in these facilities is relatively easy, as the main activity is walking. Figure 3 shows the pedestrian walking network that is considered. Notably, the two pedestrian underpasses and all access ways are included. However, the railway platforms themselves are outside the considered perimeter. Instead, all their intersections with platform access ways are represented by a centroid. It is assumed that pedestrians follow the officially signposts, which in particular implies that every OD pair is connected by exactly one route, and that there are no loops in any routes.

Regarding the time horizon, the period between 07:30 and 08:00 is considered, representing the busiest part of the morning peak hour. To quantify the day-to-day variation, the demand on multiple days is analyzed. Specifically, the morning peak hours of January 22 and 23, February 6, 27 and 28, March 5, as well as April 9, 10, 18 and 30, 2013, are considered. In the following, this ensemble of dates is referred to as the '10-day reference set'. All its elements represent either a Tuesday, Wednesday or a Thursday. The aforementioned dates have been selected by SBB based on the high punctuality of trains on these days.

A variety of information sources with different levels of aggregation and accuracy are available for the case study:

Train timetable: During the time period of interest, there are a total of 25 trains stopping at Lausanne railway station (train timetable 1). The majority of them are regional trains, subdivided into suburban (S), Regio (R) and RegioExpress (RE) trains. Additionally, there are seven express trains, classified as InterRegio (IR), InterCity (IC) and InterCity tilting (ICN) trains. These interregional trains arrive and depart in a relatively short period between 7:39 and 7:46, and 7:42 and 7:50, respectively. Not shown in table 1, for all regular trains their composition is known (SBB-Personenverkehr, 2013a,b). Besides the official schedule shown in table 1, the actual train timetable accounting for delays and changes in the track assignment is available for the 10-day reference set.

Train frequentation data[1]: To estimate passenger demand, Swiss train operators collect travel data in various ways (SBB-Personenverkehr, 2011, Olesen, 2006). Most commonly, this includes ticket sales data, infrared-based boarding and disembarkation counts, as well as within-train surveys. For the year 2010, these sources of information have been synthesized by SBB to provide an estimate of train-specific station-to-station demand. Specifically, the average number of boardings and disembarkations per year of all regular trains stopping at Lausanne is known with an accuracy of ± 3% (SBB-Personenverkehr, 2011). For use in this study, the reported figures have been increased by 15% based on the growth rate recommended by SBB (Gendre and Zulauf, 2010). These estimates are referred to as HOP-data.

Additionally, for the 10-day reference set, the mean and standard deviation of extrapolated boarding and disembarkation counts measured at train doors are available. They are referred to as FRASY-data. Their accuracy is generally lower than for HOP data, that have been synthesized from various data sources and extensively validated. FRASY-estimates for disembarkations are on average 20% higher than HOP-counts, whereas for boardings the difference amounts on average to +40%. In this case study, FRASY data is only used if no corresponding HOP

[1] Data records are not revealed due to confidentiality reasons.

Table 1: Official train timetable of Lausanne railway station between 07:30 – 08:00 (with a margin of 7 min before and after) for the period of December 11, 2011 to December 14, 2013. Columns represent the train number (train no.), associated track (#), number of rail cars (N_c), scheduled arrival time (t_a), origin of train, scheduled departure time (t_d) and destination of train. For all trains, estimates of boarding and alighting volumes are available.

Train no.	#	N_c	t_a	Origin	t_d	Destination
S21 12917	70	4			7:24	Payerne
S3 12318	8	4	7:24	Villeneuve	7:26	Allaman
S2 12217	1	4	7:26	Vallorbe	7:30	Palézieux
RE 4060	7	5	7:28	St-Maurice		
S2 12218	5	4	7:30	Palézieux	7:32	Vallorbe
S3 12317	3	4	7:33	Allaman	7:35	Villeneuve
S21 12918	70	5	7:36	Payerne		
IR 1712	6	11	7:39	Sion	7:48	Genève-Aéroport
RE 2607	1	4	7:39	Genève	7:42	Romont
IC 706	5	10	7:40	Zürich HB	7:42	Genève-Aéroport
ICN 1517	8	14			7:45	St. Gallen
IR 1407	3	9	7:42	Genève-Aéroport	7:46	Brig
IR 1710	7	7	7:42	Brig	7:45	Genève-Aéroport
IR 1606	4	9	7:43	Neuchâtel		
IR 2517	1	9	7:46	Genève-Aéroport	7:50	Luzern
RE 2710	9	8	7:49	Vevey	7:51	Genève
S 12017	5	4	7:49	Vallorbe		
S11 12820	8	5			7:55	Yverdon-les-Bains
S1 12119	3	4	7:56	Yverdon-les-Bains	8:00	Villeneuve
S4 12420	6	4	7:56	Palézieux	7:58	Allaman
RE 4024	70	5	8:00	Payerne/Romont		
S1 12120	7	4	8:02	Villeneuve	8:04	Yverdon-les-Bains
S4 12419	4	4	8:02	Allaman	8:04	Palézieux
S11 12819	3	4	8:05	Yverdon-les-Bains		
R 12014	8	6	8:06	St-Maurice	8:08	Renens

estimate is available.

An analysis of train frequentation data shows that there are more inbound than outbound passengers during the morning peak hour due to Lausanne's regional importance as a work place. Furthermore, with few exceptions express trains show a larger turnover of pedestrians than regional trains.

Historical information[1]**:** There are three sales points located in PU West, for which the average number of customer visits in 2013 is available. The corresponding number of visits during the morning peak hour is estimated by assuming that the customer frequentation is proportional to the overall occupation of the train station, i.e., that 10% of all daily sales are achieved within the morning peak hour (Lavadinho *et al.,* 2013). There are further restaurants and sales points in the train station building, represented by a generic ‘service point’ in figure 3. For this node unfortunately no customer frequentation data is available, and it is not considered in the demand analysis.

Link flow observations: For 12 links of the pedestrian walking network represented by red diamonds in figure 3, minute-by-minute count data in each direction is available. The surveyed links have been selected by the Real Estate Division of SBB, interested in knowing the afflux of pedestrians in the railway station buildings. According to SBB, the sensor technology features a very high accuracy with less than 3% error in aggregated pedestrian counts.

Trajectory recordings: For the two PUs, as well as for the interior part of platform #3/4, pedestrian trajectories are available (Lavadinho *et al.,* 2013). A silhouette-based pedestrian tracking algorithm (Alahi *et al.,* 2011) is applied to measurements of a total of about 60 visual, depth and infrared sensors, allowing to record individual pedestrian movements across space and time. The resulting trajectories are on average roughly 60% interpolated, and their largest interpolated subparts amounts on average to 20% of the total length (Babel, 2014).

4.2 Walking assignment in absence of congestion

Critical to the problem of dynamic demand estimation is the availability of representative assignment matrices, which are obtained by means of a DTA model. For the present case study, route choice is trivial, and the DTA essentially becomes a dynamic network loading model (DNL).

Two important assumptions are made regarding the DNL. First, it is supposed that walking is largely the predominant activity within the studied perimeter. This assumption is supported by various sociogeographical observations done during the morning peak hour (Lavadinho, 2012). Second, it is assumed that the level of congestion is low, i.e., that the influence of traffic on walking behavior is small. This assumption is well-founded on both sociological and quantitative observations as well (Benmoussa *et al.,* 2011). Indeed, while platforms might be highly congested, in walking facilities the level of demand-supply interaction is relatively low even during peak hours (Hänseler *et al.,* 2013). This implies that the dependency of $M(d)$ on d can be dropped, and a demand-invariant assignment results.

Based on the official dimensioning guidelines by SBB (Weidmann, 1993, Buchmüller and

Weidmann, 2008), the distribution of pedestrian walking speed is assumed to be normal

$$v \sim \mathcal{N}(1.34 \text{ m/s}, 0.34 \text{ m/s}), \tag{12}$$

where the corresponding probability density and cumulated distribution functions are denoted by $f_v(v)$ and $F_v(v)$, respectively. Expression 12 has been obtained indirectly from observations of travel times and traveled distances on even walking areas. In this work, it will be used in the same context, i.e., to estimate the distribution of travel times. On inclined walking areas or stairways, the velocity of pedestrians deviates from that observed on even areas. According to Weidmann (1993), the horizontal speed on stairways averages to 0.61 m/s for pedestrians walking upward, whereas an average of 0.694 m/s is reported for pedestrians walking downward. For ramps with an inclination of 15%, the corresponding velocities are estimated at 1.07 and 1.40 m/s. It is assumed that the standard deviation of the distribution changes proportionally with the change in average speed.

Based on the speed distribution 12, the assignment mappings can be derived for various demand indicators including link flows and area occupations. Let the distance along a route ρ up to the beginning of edge λ be denoted by ℓ_ρ^λ. Furthermore, let the departure times of pedestrians within a time interval be distributed uniformly, i.e., the distribution of continuous departure time k associated with demand $d_{\rho,\kappa}$ for a route ρ and a time interval κ is given by

$$h_\kappa(k) = \begin{cases} \frac{1}{\Delta t} & \text{if } k \in \kappa, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Assuming that each pedestrian is walking at a constant speed, the probability for a person on route ρ that departed during time interval κ to arrive on link λ during time interval τ is given by

$$\begin{aligned} \hat{a}_{(\lambda,\tau),(\rho,\kappa)} &= \widehat{\Pr}(k \in \kappa, t \in \tau | \rho, \lambda) \\ &= \widehat{\Pr}\left(k \in \kappa, v \in \left[\frac{\ell_\rho^\lambda}{t_\tau^+ - k}, \frac{\ell_\rho^\lambda}{t_\tau^- - k}\right]\right), \end{aligned} \tag{14}$$

where k and t represent the (continuous) departure and arrival time, respectively. For the most common case that $\ell_\rho^\lambda > 0$ and $\tau > \kappa$, we obtain

$$\begin{aligned} \hat{a}_{(\lambda,\tau),(\rho,\kappa)} &= \int_{k=k^-}^{k^+} \int_{v=\ell_\rho^\lambda/(t_\tau^+ - k)}^{\ell_\rho^\lambda/(t_\tau^- - k)} f_v(v) g_\kappa(k) \, \mathrm{d}v \, \mathrm{d}k \\ &= \frac{1}{\Delta t} \int_{k=k^-}^{k^+} F_v\left(\frac{\ell_\rho^\lambda}{t_\tau^- - k}\right) - F_v\left(\frac{\ell_\rho^\lambda}{t_\tau^+ - k}\right) \mathrm{d}k. \end{aligned} \tag{15}$$

Similarly, if $\ell_\rho^\lambda > 0$ and $\kappa = \tau$, we obtain

$$\begin{aligned}
\hat{a}_{(\lambda,\tau),(\rho,\tau)} &= 1 - \Pr\left(k \in \tau, t \notin \tau | \rho, \lambda\right) \\
&= 1 - \Pr\left(k \in \tau, v \in \left[0, \frac{\ell_\rho^\lambda}{t_\tau^+ - k}\right]\right) \\
&= 1 - \frac{1}{\Delta t}\int_{k=t_\tau^-}^{t_\tau^+} F_v\left(\frac{\ell_\rho^\lambda}{t_\tau^+ - k}\right) - F_v(0)\ \mathrm{d}k.
\end{aligned} \tag{16}$$

Thus, the assignment mapping corresponding to equation 4 is given by

$$\hat{a}_{(\lambda,\tau),(\rho,\kappa)} = \begin{cases} 1 & \text{if } \ell_\rho^\lambda = 0, \kappa = \tau, \\ 0 & \text{if } \ell_\rho^\lambda = 0, \kappa < \tau, \\ \text{eq. 16} & \text{if } \ell_\rho^\lambda > 0, \kappa = \tau, \\ \text{eq. 15} & \text{if } \ell_\rho^\lambda > 0, \kappa < \tau. \end{cases} \tag{17}$$

The assignment mapping for area occupation can be derived accordingly. Let us consider an area α, and let us assume that each route enters and leaves area α at most once. Let v be the constant, individual speed of a person traveling along route ρ, $\ell_{in}^{\rho,\alpha}$ the distance along the route ρ to the entrance of area α and $\ell_{out}^{\rho,\alpha}$ the corresponding distance to the exit. Consequently, $t_{in} = \ell_{in}^{\rho,\alpha}/v$ is the time after departure at which a person with speed v enters area α and $t_{out} = \ell_{out}^{\rho,\alpha}/v$ the corresponding time at which he exits. If a route ρ does not cross area α, then $\ell_{in}^{\rho,\alpha} = \infty$. If we consider a time interval $[t^-, t^+]$ after departure, the expected sojourn time for this person with constant speed v inside the area α within the interval is given by

$$\sigma(v, \ell_{in}^{\rho,\alpha}, \ell_{out}^{\rho,\alpha}, t^-, t^+) = \begin{cases} t^+ - \ell_{in}^{\rho,\alpha}/v & \text{if } t^- \le \ell_{in}^{\rho,\alpha}/v \le t^+ \le \ell_{out}^{\rho,\alpha}/v, \\ \ell_{out}^{\rho,\alpha}/v - t^- & \text{if } \ell_{in}^{\rho,\alpha}/v \le t^- \le \ell_{out}^{\rho,\alpha}/v \le t^+, \\ t^+ - t^- & \text{if } \ell_{in}^{\rho,\alpha}/v \le t^- \le t^+ \le \ell_{out}^{\rho,\alpha}/v, \\ (\ell_{out}^{\rho,\alpha} - \ell_{in}^{\rho,\alpha})/v & \text{if } t^- \le \ell_{in}^{\rho,\alpha}/v \le \ell_{out}^{\rho,\alpha}/v \le t^+, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

In equation 18, the first line corresponds to the case where a person reaches the area within the time interval, but does not exit it. The second line is the inverse case. The third line represents the case where a person stays within the area during the full time period. Finally, the fourth line represents the case where a pedestrian enters and leaves the area during the period of interest, and the fifth case the situation where a pedestrian is not present in area α during the time interval at all.

Using equations 5 and 18, the 'occupation contribution' of a pedestrian traveling along route ρ

with departure time interval κ in area α during time interval τ is given by

$$\begin{aligned} \hat{s}_{(\alpha,\tau),(\rho,\kappa)} &= \int_{t=t_\kappa^-}^{t_\kappa^+} \int_{v=0}^{\infty} \frac{\sigma(v, \ell_{in}^{\rho,\alpha}, \ell_{out}^{\rho,\alpha}, t_\tau^- - t, t_\tau^+ - t)}{\Delta t} f_v(v) h_\kappa(t) \; \mathrm{d}v \; \mathrm{d}t \\ &= \frac{1}{\Delta t^2} \int_{v=0}^{\infty} f_v(v) \int_{t=t_\kappa^-}^{t_\kappa^+} \sigma(v, \ell_{in}^{\rho,\alpha}, \ell_{out}^{\rho,\alpha}, t_\tau^- - t, t_\tau^+ - t) \; \mathrm{d}t \; \mathrm{d}v. \end{aligned} \tag{19}$$

For an efficient implementation, we note that the assignment mappings 17 and 19 are time-invariant, i.e., for $\delta = \tau - \kappa$ it holds that

$$\hat{a}_{(\lambda,\tau),(\rho,\kappa)} = \hat{a}'_{\lambda,\rho,\delta} \text{ and } \hat{s}_{(\alpha,\tau),(\rho,\kappa)} = \hat{s}'_{\alpha,\rho,\delta}. \tag{20}$$

To further reduce the computational cost, a maximum travel time TT_{max} is defined. If $\delta \geq TT_{max}$, it is assumed that $\hat{a}'_{\lambda,\rho,\delta} = 0 \; \forall \lambda, \rho$ and $\hat{s}'_{\alpha,\rho,\delta} = 0 \; \forall \alpha, \rho$. The threshold TT_{max} is chosen large enough such that the error incurred by this approximation is negligible.

4.3 Specification of model for train-induced arrival flows

Molyneaux *et al.* (2014) investigate train-induced arrival flows on platform access ways in Lausanne railway station. For each link λ marked with a yellow asterisk in figure 3, a relationship of the form of equation 7 is established. Let C_λ denote the capacity flow rate of link λ, and s_λ the corresponding dead time representing the time lag between the train arrival and the onset of the flow. According to Molyneaux *et al.* (2014), the time-dependent arrival flow associated with train m is given by

$$y_\lambda^{arr}(t; Q_m^{al}, t_m^{arr}, C_\lambda, s_\lambda) = \begin{cases} C_\lambda & t \in \left(t_m^{arr} + s_\lambda, t_m^{arr} + s_\lambda + Q_m^{al}/C_\lambda\right), \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

The alighting volume Q_m^{al} is estimated from HOP or FRASY data, representing a normally distributed random variable with a standard deviation equal to 19.2% of its mean. This specification has been derived from an analysis of FRASY data and is in good agreement with results from the analysis of pedestrian trajectory data (Molyneaux *et al.,* 2014). The dead time s_λ and the capacity flow rate C_λ are also stochastic variables. Moreover, on some links the capacity flow rate C_λ depends on the alighting volume Q_m^{al}.

For a detailed description of the model specification and its parameters, see Molyneaux *et al.* (2014). For the sake of completeness, it is noted that a model of the form of equation 8 has not been developed yet for Lausanne railway station due to a lack of appropriate data.

4.4 A priori route split fractions at origins

As mentioned previously, four visitor types are prevalent in Lausanne railway station. If a pedestrian enters the railway station through a platform, he is either an inbound or a transfer passenger. Otherwise, he is either a local user or an outbound passenger. From sociogeographical observations, the shares of each visitor type are approximatively known. This information can be incorporated in the demand estimation process.

To do so, let us first divide the set of centroids into two subsets C^p and C^{np} based on the notion of a node type $u = \{p, np\}$. The subset C^p contains all nodes associated with platforms, and the subset C^{np} all the remaining ones, i.e., the 'non-platforms'. The mentioned sociogeographical observations can then be formalized as follows. Let the fraction of disembarkations that is accounted for by inbound passengers be represented by random variable $\beta_{p\to np}$, and the corresponding fraction of transfer passengers by $\beta_{p\to p} = (1 - \beta_{p\to np})$. Similarly, let the fraction of outbound passengers among the users entering the train station through a non-platform be given by the random variable $\beta_{np\to p}$, and the corresponding fraction of local users by $\beta_{np\to np} = (1 - \beta_{np\to p})$.

In the case of Lausanne railway station, an estimate of time-aggregated destination flows is available for all centroids. For platform centroids, these can be estimated from the cumulated sum of boarding volumes and the assignment of these to the platform access ways (see Molyneaux *et al.,* 2014, for details). For non-platforms such as shops and exits, sales data and pedestrian counts are available, from which a static estimate can be obtained. Let the a priori destination flow in centroid v cumulated over the time period $\mathcal{T}$ be denoted by the random variable b_v.

The cumulated destination flows may be used to estimate a priori route split fractions. Let the route split fraction r_ρ be defined as the average fraction of pedestrians emanating from note v_o^ρ that is associated with route ρ during the time horizon $\mathcal{T}$. Moreover, let the ensemble of routes starting in node $v \in C$ be represented by $\mathcal{R}_v$. We may distinguish between routes leading to platform and non-platform nodes, denoted by $\mathcal{R}_v^p$ and $\mathcal{R}_v^{np}$, respectively. Assuming that for a given destination type u_d, the split ratio of a route $\rho \in \mathcal{R}_{v_o^\rho}^{u_d}$ emanating from node $v_o^\rho \in C^{u_o}$ is proportional to the destination flow $b_{v_d^\rho}$ in v_d^ρ, we have

$$\hat{r}_\rho = \beta_{u_o^\rho \to u_d^\rho} \frac{b_{v_d^\rho}}{\sum_{\rho' \in \mathcal{R}_{v_o^\rho}^{u_d}} b_{v_d^{\rho'}}}. \tag{22}$$

Clearly, it holds that $\sum_{\rho \in \mathcal{R}_v} \hat{r}_\rho = 1$. Equation 22 can be interpreted as a hierarchical route choice model that consists of two terms representing the destination type and node choice. By reordering terms and defining a corresponding a priori route split fraction matrix $\hat{M}_r$, equation 22 can be represented in matrix form as

$$\hat{M}_r d + \epsilon_r = 0, \tag{23}$$

which complies with the standard assignment form shown in equation 2.

For Lausanne railway station, during the morning peak period a fraction of $\beta_{p \to np} = 91.35\% \pm 4.55\%$ of disembarking passengers in Lausanne represent inbound passengers (Anken *et al.*, 2012). Inversely, the percentage of outbound passengers among pedestrians entering the train station from the city amounts to about $\beta_{np \to p} \approx 95\%$, as reported independently by two studies (Benmoussa *et al.*, 2011, Lavadinho *et al.*, 2013).

4.5 Estimation and validation of demand

The previously presented information sources are used for the estimation and validation of demand in various ways. A priori information of demand is first obtained by estimating train-induced arrival flows from the train timetable and alighting volumes, as well as from a choice model that provides route split fractions for all centroids. This a priori information is combined with flow observations that are available for several links in the walking network. Since the resulting estimation problem is intrinsically under-determined, a zero prior with a suitable small weight is additionally used to guides the selection of solutions towards one with maximum entropy (Cascetta *et al.*, 1993). For validation, trajectory recordings are used to compute the occupation and the OD demand in the two pedestrian underpasses, which can be compared to model estimates. An overview of the information flow in the demand estimation process is provided by figure 4.

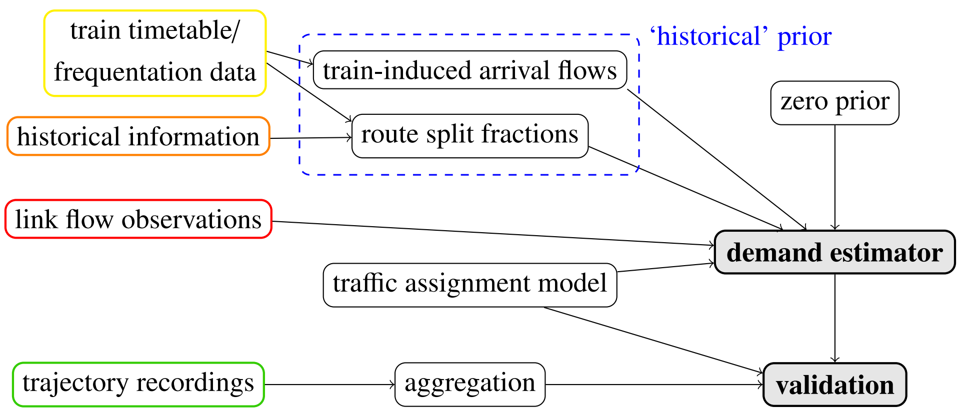

Figure 4: Scheme of the demand estimation framework applied to a case study of Lausanne railway station. The use of each data source for the generation of a 'historical prior', for the estimation and for validation is illustrated. The color scheme corresponds to figure 3.

The present case study aims at estimating the demand during the peak period between 07:30 and 08:00. To account for artificial transients in the demand estimates during a potential 'heat-up' of the estimation, the computations include an additional 7 minutes both at the beginning and the end of the period. The temporal aggregation is one minute. This choice is constrained by the aggregation of the link flow counts, which do not allow for a finer resolution without incurring a substantial under-determination. In total, 44 time intervals of 60s are thus considered.

As a distance measure for any indicator q, the square of the Euclidian norm with weight μ_q is used, i.e.,

$$\text{dist}(\hat{q}, \hat{M}_q(d)d) = \mu_q \left\|\hat{q} - \hat{M}_q(d)d\right\|_2^2, \tag{24}$$

where a meaningful choice of the weights is $\mu_q = 1/\operatorname{Var}(\epsilon_q)$. In practice, these weights need to be estimated, since the distribution of the error terms is unknown.

Since the a priori estimates of train-induced flows and route split fractions depend on stochastic parameters, the demand estimate d^* defined in equation 11 represents a random variable as well. To estimate its distribution, Monte Carlo integration may be used. Let ω be the vector containing all stochastic parameters, and let its distribution density be denoted by $p(\omega)$. If $\omega_1, \ldots, \omega_N$ represent independent draws from $p(\omega)$, and if $d^*(Q, \mathcal{M}|\omega)$ represents the demand estimate resulting for a given draw ω, the expected value of d^* is given by

$$\mathbb{E}_{p(\omega)}\left[d^*(Q, \mathcal{M})\right] \approx \frac{1}{N}\sum_{i=1}^{N} d^*(Q, \mathcal{M}|\omega_i). \tag{25}$$

Other statistical measures, such as for instance the variance-covariance matrix, can be derived accordingly.

5 Conclusions

During peak hours, rail access installations in large train stations often reach capacity and may reduce the performance of a transportation system as a whole. To optimize their design and operation, there is thus a general need to better understand pedestrian behavior in railway stations. An increasing effort is made towards this end both by operators of railway networks and academia. However, most researchers and practitioners concentrate on investigating the interaction between pedestrian demand and infrastructure, whereas the estimation of pedestrian demand as such has received relatively little attention so far.

In this study, a framework for the time-dependent estimation of pedestrian origin-destination demand within a train station has been presented. Besides direct and indirect demand indicators such as flow counts or sales data, the train timetable is explicitly taken into account. This may be achieved by establishing an empirical relation between the departure of a train and the accumulation of prospective train passengers on platforms, as well as the arrival of a train and the subsequent flow of alighting passengers on platform exit ways. The formulation of the framework is such that it can be applied to various types of railway stations and may be used with different data sources.

Besides a detailed theoretical consideration of the framework, a case study of Lausanne railway station has been presented, and various methodological and practical challenges have been discussed.

Acknowledgement

Financial support by SNF grant #200021-141099 'Pedestrian dynamics: flows and behavior' and by SBB-CFF-FFS in the framework of 'PedFlux', as well as support by Quentin Mazars-Simon and Eduard Rojas in implementing the case study of Lausanne railway station is gratefully acknowledged.

6 References

Alahi, A., L. Bagnato, D. Chanel and A. Alahi (2013) Technical report for SBB network of sensors, *Technical Report*, VisioSafe.

Alahi, A., L. Jacques, Y. Boursier and P. Vandergheynst (2011) Sparsity driven people localization with a heterogeneous network of cameras, *Journal of Mathematical Imaging and Vision*, **41** (1-2) 39–58.

Amacker, K. (2012) SBB Facts and Figures, *Technical Report*, Swiss Federal Railways (SBB-CFF-FFS).

Anken, N., F. Hänseler and M. Bierlaire (2012) Flux piétonniers dans la gare de Lausanne: Vers l'estimation d'une matrice OD à l'aide des extrapolations voyageurs des CFF, *Technical Report*, EPFL.

Babel, H. (2014) Exploring pedestrian mobility using video tracking data in Lausanne train station, Semester thesis, Ecole Polytechnique Fédérale de Lausanne.

Benmoussa, M., F. Ducommun, A. Khalfi, M. Kharouf, A. Koymans, M. Nguyen, A. Raies,

M. Vidaud and C. Birchler (2011) Analyse des flux piétonniers en gare de Lausanne, *Technical Report*, Ecole Polytechnique Fédérale de Lausanne.

Bera, S. and K. Rao (2011) Estimation of origin-destination matrix from traffic counts: The state of the art, *European Transport*, **49**, 3–23.

Buchmüller, S. and U. Weidmann (2008) *Handbuch zur Anordnung und Dimensionierung von Fussgängeranlagen in Bahnhöfen*.

Cascetta, E., D. Inaudi and G. Marquis (1993) Dynamic estimators of origin-destination matrices using traffic counts, *Transportation Science*, **27** (4) 363–373.

Cheung, C. and W. Lam (1998) Pedestrian route choices between escalator and stairway in MTR stations, *Journal of Transportation Engineering*, **124** (3) 277–285.

Daamen, W. (2004) Modelling passenger flows in public transport facilities, Ph.D. Thesis, Delft University of Technology.

Daamen, W., P. H. Bovy and S. P. Hoogendoorn (2005) Influence of changes in level on passenger route choice in railway stations, *Transportation Research Record: Journal of the Transportation Research Board*, **1930** (1) 12–20.

Daamen, W., Y. Lee and P. Wiggenraad (2008) Boarding and alighting experiments: Overview of setup and performance and some preliminary results, *Transportation Research Record: Journal of the Transportation Research Board*, **2042** (1) 71–81.

Daly, P., F. McGrath and T. Annesley (1991) Pedestrian speed/flow relationships for underground stations, *Traffic Engineering & Control*, **32** (2) 75–78.

Danalet, A., B. Farooq and M. Bierlaire (2014) A Bayesian approach to detect pedestrian destination-sequences from WiFi signatures, *Transportation Research Part C: Emerging Technologies*. Forthcoming.

Davidich, M., F. Geiss, H. Mayer, A. Pfaffinger and C. Royer (2013) Waiting zones for realistic modelling of pedestrian dynamics: A case study using two major german railway stations as examples, *Transportation Research Part C: Emerging Technologies*.

Edie, L. C. (1963) Discussion of traffic stream measurements and definitions, *Proceedings of the Second International Symposium on the Theory of Traffic Flow*, 139–154.

Gendre, G. and C. Zulauf (2010) Gare de Lausanne: Analyse des flux piétonniers, *Technical Report*, Swiss Federal Railways (SBB-CFF-FFS).

Hanisch, A., J. Tolujew, K. Richter and T. Schulze (2003) Online simulation of pedestrian flow in public buildings, paper presented at the *Winter Simulation Conference*, vol. 2, 1635–1641.

Hänseler, F., N. Molyneaux, M. Thémans and M. Bierlaire (2013) Pedestrian strategies within railway stations: Analysis and modeling of pedestrian flows (PedFlux Mid-Term Report), *Technical Report*, EPFL.

Hermant, L. (2012) Video data collection method for pedestrian movement variables & development of a pedestrian spatial parameters simulation model for railway station environments, Ph.D. Thesis, Stellenbosch University.

Hermant, L., M. De Gersigny, R. Hermann and R. Ahuja (2010) Applying microscopic pedestrian simulation to the design assessment of various railway stations in south africa, paper presented at the *Proceedings of the 29th Southern African Transport Conference (SATC 2010)*, vol. 16, 19.

Hoogendoorn, S. and P. Bovy (2004) Pedestrian route-choice and activity scheduling theory and models, *Transportation Research Part B: Methodological*, **38** (2) 169–190.

Hoogendoorn, S. and W. Daamen (2004) Design assessment of Lisbon transfer stations using microscopic pedestrian simulation, paper presented at the *Computers in railways IX (Congress Proceedings of CompRail 2004)*, 135–147.

Jiang, C., Y. Deng, C. Hu, H. Ding and W. K. Chow (2009) Crowding in platform staircases of a subway station in China during rush hours, *Safety Science*, **47** (7) 931–938.

Kaakai, F., S. Hayat and A. El Moudni (2007) A hybrid Petri nets-based simulation model for evaluating the design of railway transit stations, *Simulation Modelling Practice and Theory*, **15** (8) 935–969.

Lam, W. and C. Cheung (2000) Pedestrian speed/flow relationships for walking facilities in Hong Kong, *Journal of Transportation Engineering*, **126** (4) 343–349.

Lam, W., C. Cheung and C. Lam (1999) A study of crowding effects at the Hong Kong light rail transit stations, *Transportation Research Part A: Policy and Practice*, **33** (5) 401–415.

Lavadinho, S. (2012) Compréhension fine des stratégies piétonnières en gare de Lausanne, *Technical Report*, Ecole Polytechnique Fédérale de Lausanne.

Lavadinho, S., A. Alahi and L. Bagnato (2013) Analysis of Pedestrian Flows: Underground pedestrian walkways of Lausanne train station, *Technical Report*, VisioSafe SA.

Lee, J., W. Lam and S. Wong (2001) Pedestrian simulation model for Hong Kong underground stations, paper presented at the *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*, 554–558.

Molyneaux, N., F. Hänseler and M. Bierlaire (2014) Modeling of train-induced pedestrian flows in railway stations, *STRC Proceedings*.

Olesen, A. (2006) HOP(P) Schwiiz, *Info Retica*, **2**, 34–35.

Pettersson, P. (2011) Passenger waiting strategies on railway platforms: Effects of information and platform facilities, Master's thesis, KTH.

Rindsfüser, G. and F. Klügl (2007) Agent-based pedestrian simulation: A case study of Bern Railway Station, *The Planning Review*, **170**, 9–18.

SBB-Infrastruktur (2013) Le processus de l'élaboration de l'horaire, *Technical Report*, Swiss Federal Railways (SBB-CFF-FFS).

SBB-Personenverkehr (2011) SBB Clearing Extrapolation Voyageurs HOP, *Technical Report*, Swiss Federal Railways (SBB-CFF-FFS).

SBB-Personenverkehr (2013a) Zugbildungsplan Fernverkehr, *Technical Report*, Swiss Federal Railways (SBB-CFF-FFS).

SBB-Personenverkehr (2013b) Zugbildungsplan Regionalverkehr West, *Technical Report*, Swiss Federal Railways (SBB-CFF-FFS).

Schneider, J. (2012) Network condition report 2012, *Technical Report*, Swiss Federal Railways (SBB-CFF-FFS).

Tolujew, J. and F. Alcalá (2004) A mesoscopic approach to modeling and simulation of pedestrian traffic flows, paper presented at the *Proceedings of 18th European Simulation Multiconference*.

Weidmann, U. (1993) *Transporttechnik der Fussgänger*, Institute for Transport Planning and Systems, ETH Zürich.

Xu, X., J. Liu, H. Li and J. Hu (2014) Analysis of subway station capacity with the use of queueing theory, *Transportation Research Part C: Emerging Technologies*, **38**, 28–43.

Zhang, Q., B. Han and D. Li (2008) Modeling and simulation of passenger alighting and boarding movement in Beijing metro stations, *Transportation Research Part C: Emerging Technologies*, **16** (5) 635–649.