Preliminary ideas for dynamic estimation of pedestrian origin-destination demand within train stations

Flurin Hänseler
Bilal Farooq
Michel Bierlaire

STRC 2012 **May 2012**

STRC 2012

Transport and Mobility Lab
School of Architecture, Civil and Environmental Engineering
École Polytechnique Fédérale de Lausanne (EPFL)

Preliminary ideas for dynamic estimation of pedestrian origin-destination demand within train stations

Flurin Hänseler
flurin.haenseler@epfl.ch

Bilal Farooq
bilal.farooq@epfl.ch

Michel Bierlaire
michel.bierlaire@epfl.ch

May 2012

Abstract

This article provides preliminary ideas for the dynamic estimation of pedestrian origin-destination (OD) demand within train stations. In particular, a methodology is outlined that can predict OD demand as function of the train time table and train track assignment. At the example of a simple case study, the methodology is concretized and elaborated.

Keywords

pedestrian flows, origin-destination demand

1 Introduction

In the last decades, pedestrian flows in transportation hubs like train stations or airports have increased substantially. At peak hours, pedestrian facilities get congested easily, which has a negative influence on overall customer satisfaction. In other words, pedestrian flows in transfer hubs have become an important factor for a transportation network as a whole. With the anticipated increase in travel demand, phenomena related to pedestrian flows such as human jams will gain even more importance in the future.

Pedestrian flows have a considerable influence on performance, perceived comfort and safety of a transportation hub. For instance, in order to guarantee timetable stability, transit times need to be chosen such that passengers have enough time to catch their connection. In general, the more fluid pedestrian flows are, the shorter travel times become, and the higher the performance of a transfer station. The more efficient flows are and the less crowded the pedestrian facilities, the more comfortable people feel. Last but not least, in case of an emergency, the right management of pedestrian flows is key in avoiding a stampede.

In order to get a better understanding of phenomena such as congestion in pedestrian facilities, there is a general need to analyze and model pedestrian flows. Pedestrian models allow to simulate complex travel demand scenarios and to quantitatively evaluate the suitability of an infrastructure layout using level of service indicators. In the future, flow models will allow optimizing the design of pedestrian facilities as well as their operation. For example, it could be studied how the train time table or train track assignment can be changed in order to minimize human jams, or it could be investigated if active flow controlling by signalization can alleviate congestion.

2 Modeling pedestrian flows in transfer stations

Pedestrian flow models for transfer stations need to take different levels of behavior into account. In a nutshell, pedestrian origin-destination demand needs to be estimated, each pedestrian needs to be assigned a route that takes him from his origin to his desired destination, and then a walking model has to be used to derive the actual trajectories of pedestrians. In a similar context, Hoogendoorn *et al.* referred to these three phases as the strategical, tactical, and operational level. Especially if congestion is present, the three levels are inherently coupled, which makes the modeling process very interesting. For example, people might change their minds about going to a certain location if all the links leading there are overcrowded, or at least chose a different route, and so on.

While a lot has been done on solving similar problems for car traffic networks, pedestrian flow modeling is still in its infancy. In particular, estimation of pedestrian OD demand has not received much scientific attention so far, despite its increasing importance. This is where this report would like to make a contribution for the specific case of a train station.

For a detailed discussion of the importance to study pedestrian flows in train stations, as well as for an excellent comparison between pedestrian and car traffic, the interested reader is referred to Daamen (1999, 2002, 2004).

3 Estimating origin-destination demand in a train station

In the following, a methodology for the estimation of pedestrian origin-destination demand within train stations is outlined. A general framework is developed, and then applied to a case study.

Consider the state variable $x_{i,j,t}$, which denotes the pedestrian demand rate from node i to node j at departure time t. A node at which trips originate and/or terminate is called a centroid (Cascetta and Nguyen, 1988), which are indexed by $r = 1, \ldots, R$. Time is discretized in intervals of uniform length indexed by $t = 1, \ldots, T$. A suitable interval length is expected to be in the order of a few minutes. The discrete random variable $y_{i,j,t}$ represents the number of time intervals required to reach destination j when leaving origin i during time interval t.

The origin flow of centroid i is defined as

$$f_{i,t} = \sum_{j=1}^{R} x_{i,j,t}. \tag{1}$$

which is the total number of pedestrians leaving origin i during time interval t. Similarly, the destination flow

$$g_{j,t} = \sum_{k=1}^{t} \sum_{i=1}^{R} x_{i,j,k} P(i,j,k,t) = \sum_{k=1}^{t} \sum_{i=1}^{R} x_{i,j,k} \Pr(y_{i,j,k} = t-k), \tag{2}$$

is the total number of pedestrians reaching destination j during time interval t where

$$P(i,j,k,t) = \Pr(y_{i,j,k} = t-k) \tag{3}$$

is the transition probability that a pedestrian leaving origin i during time interval k reaches destination j during time interval t, which is the probability that the corresponding travel time represents $t-k$ time intervals.

For nodes where count data is available, measurement equations for the origin and destination flows can be expressed as

$$\widehat{f}_{i,t} = f_{i,t} + \xi_{i,t} \quad \forall i \in F, t \tag{4}$$

and

$$\widehat{g}_{j,t} = g_{j,t} + \nu_{j,t} \quad \forall j \in G, t \tag{5}$$

respectively. F and G are the sets of centroids for which outgoing and incoming flow counts are available, respectively. Such nodes might include for example the main entrances of a train station. The variables $\xi_{i,t}$, $\nu_{j,t}$ represent random terms accounting for dynamics that are not covered by the deterministic part of the model.

For individual platforms in train stations, pedestrian count data is often unavailable. In contrast to that, railway operators usually have precise data on how many passengers get on and off a given train. This information can be used together with the train time table and track assignment to estimate flows towards and away from platforms. In the following, a methodology is developed for this purpose.

For each centroid j representing a platform, the total number of trains using an adjacent track is denoted by N_j. For such a train z, the time interval associated with its arrival, the interval associated with its departure, the passenger capacity, as well as the number of people getting off and on (relative to the capacity) are given by $a_{j,z}$, $b_{j,z}$, $q_{j,z}$, as well as $o_{j,z}$ and $p_{j,z}$, respectively. The number of disembarking and boarding passengers can thus be expressed as

$$\phi_{j,z} = q_{j,z} o_{j,z} + \varepsilon_{j,z}$$

and

$$\pi_{j,z} = q_{j,z} p_{j,z} + \eta_{j,z}$$

respectively. The variables $\varepsilon_{j,z}$ and $\eta_{j,z}$ are random variables taking the fluctuations in passenger number into account. Their distribution is known based on historical data.

Findings of a european railway operator[1] suggest that the arrival pattern on a platform of pedestrians about to board a train follows a beta distribution. Specifically, this pattern can be expressed as

$$\tilde{B}_p(\tilde{t}; \tilde{\gamma}, \tilde{\delta}, \tilde{t}_p) = \frac{(\tilde{t} - \tilde{t}_p + 1)^{\tilde{\gamma}-1} (\tilde{t}_p - \tilde{t})^{\tilde{\delta}-1}}{\int_0^1 u^{\tilde{\gamma}-1} (1-u)^{\tilde{\delta}-1} \, du} \tag{6}$$

[1] We are not allowed to disclose the name.

where $\tilde{t}$ denotes continuous time, $\tilde{t}_p$ the departure time, and $\tilde{\gamma}$ and $\tilde{\delta}$ are two positive shape parameters. Figure 1 shows both the rate of arrival as well as the total number of pedestrians on the platform as function of time.

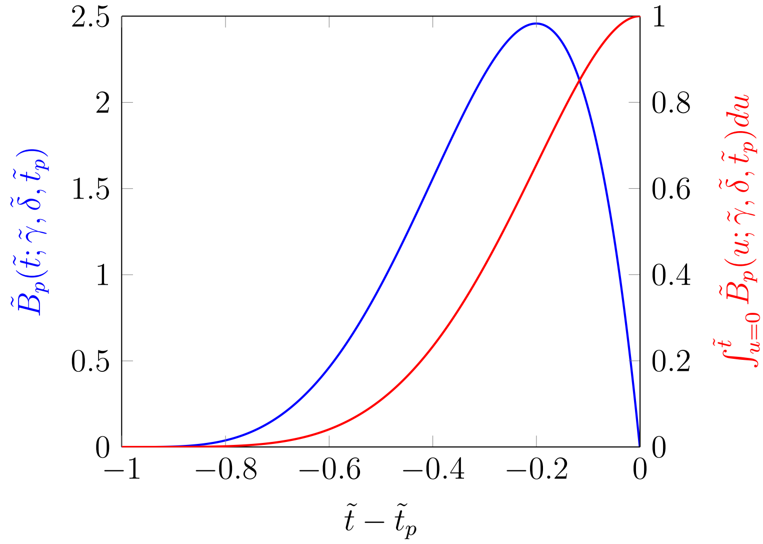

Figure 1: The rate of arrival (blue) as well as the total number of pedestrians (red) awaiting to board a train as a function of time. For this example, $\tilde{\gamma}=5$ and $\tilde{\delta}=2$ was chosen.

The pedestrian flow caused by passengers disembarking from a recently arrived train can be described analogically by

$$\tilde{B}_o(\tilde{t};\tilde{\alpha},\tilde{\beta},\tilde{t}_o) \sim \tilde{B}_p(-\tilde{t};\tilde{\gamma},\tilde{\delta},-\tilde{t}_p)$$

which is qualitatively the same function as above, just flipped in time. Besides, the scaling of the functions is usually different, as pedestrian waves induced by incoming trains are typically more intense than those of outgoing trains, where people start pouring in already several minutes before departure.

Let $\{\alpha,\beta\}_{j,z}$ and $\{\gamma,\delta\}_{j,z}$ be the shape parameters for the pedestrian flows induced by the departure and arrival of train z on platform j. Furthermore, let $B_o\left(t;\alpha_{j,z},\beta_{j,z},a_{j,z}\right)$ and $B_p\left(t;\gamma_{j,z},\delta_{j,z},b_{j,z}\right)$ be the discrete counterparts of $\tilde{B}_o$ and $\tilde{B}_p$, i.e. the discrete, deterministic flow patterns of pedestrians. For a given time interval t, structural equations for centroids representing train platforms can be formulated as follows

$$d_{i,t} = \sum_{z=1}^{N_i} \phi_{i,z} B_o\left(t;\alpha_{i,z},\beta_{i,z},a_{i,z}\right) \tag{7}$$

$$e_{j,t} = \sum_{z=1}^{N_j} \pi_{j,z} B_p\left(t;\gamma_{j,z},\delta_{j,z},b_{j,z}\right) \tag{8}$$

where the first equation describes the flow pattern of pedestrians disembarking and streaming into the train station, and the second equation the flow pattern of pedestrians showing up on the platform in order to board a train. It is important to note that in this formulation, B_o and B_p represent concrete values since both functions are evaluated at a given time interval t and all parameters are known. This assumption is not necessary and could be relaxed. However, here it is a suitable simplification that allows formulating the equations without introducing further variables.

The above framework can finally be completed by writing two measurement equations

$$\widehat{d}_{i,t} = f_{i,t} + \zeta_{i,t} \qquad \forall\, i \in I, t \tag{9}$$

$$\widehat{e}_{j,t} = g_{j,t} + \lambda_{j,t} \qquad \forall\, j \in J, t \tag{10}$$

I and J represent the centroids for which the origin and destination flows are estimated using information related to trains arriving on and leaving form the corresponding platforms, respectively. $\zeta_{i,t}$ and $\lambda_{j,t}$ represent random variables. This formulation of the measurement equations represents a rather simple approach but still allows to convey the main ideas of the estimation methodology in a clear way. In future work, the present framework can be extended and completed in many ways.

4 Case Study

In the following, the above concept is detailed at the example of a simple, widely-spread train station layout (figure 2), with Renens being an example particularly close to EPFL. Several platforms are aligned along a single underpass, which also connects two opposite urban districts.

This example features five centroids, two for the interface with the city, and three for platforms serving five tracks in total. The pedestrian network for the general case with R centroids is shown in figure 3. Nodes 1 and R can be seen as the neighboring districts, from and to which people go passing the railway station, and nodes $2 \ldots (R-1)$ represent platforms. In addition, $(R-2)$ intersection nodes connect these centroids. As no new people enter or leave the train station through these nodes, they fulfill flow conservation.

While intersection nodes do not appear in the OD matrix, they are indispensable for the route assignment. Together with these nodes, the link flows $\ell_{i,j,t}$ for pedestrians leaving from node i during time interval t towards node j are introduced. For the intersection node $(j + R - 1)$, the

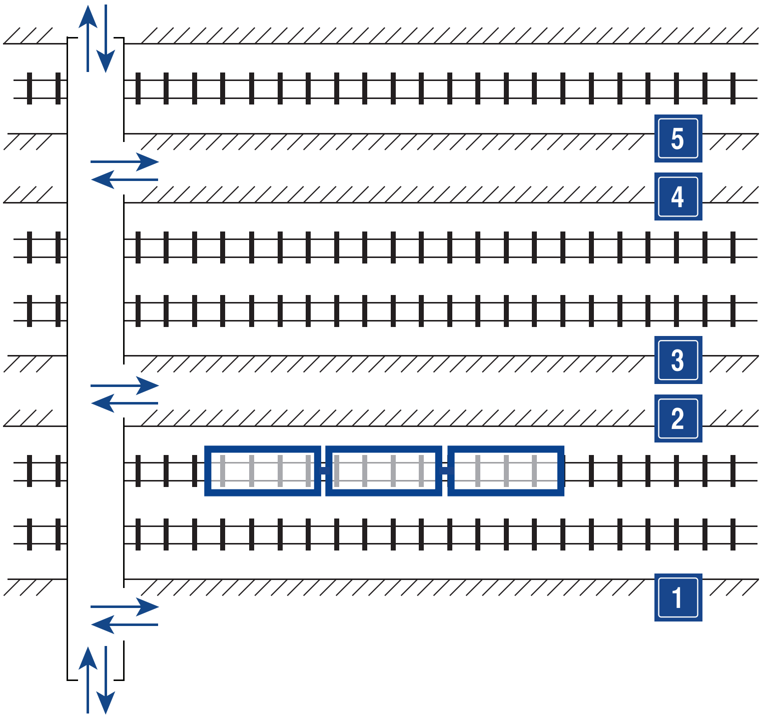

Figure 2: A typical layout for a small- to medium-size train station in Switzerland. A single underpass connects two opposite neighborhoods and the platforms of the train station. Arrows denote pedestrian in- and outflows.

corresponding link flows can be expressed as

$$\begin{aligned}
\ell_{j+R-2,j+R-1,t} &= \textstyle\sum_{k=1}^{t} \sum_{m=1}^{j-1} \sum_{n=j}^{R} x_{m,n,k} P(m, j+R-2, k, t) \\
\ell_{j,j+R-1,t} &= f_{j,t} \\
\ell_{j+R,j+R-1,t} &= \textstyle\sum_{k=1}^{t} \sum_{m=j+1}^{R} \sum_{n=1}^{j} x_{m,n,k} P(m, j+R, k, t)
\end{aligned}$$

for incoming flows and

$$\begin{aligned}
\ell_{j+R-1,j+R-2,t} &= \textstyle\sum_{k=1}^{t} \sum_{m=j}^{R} \sum_{n=1}^{j-1} x_{m,n,k} P(m, j+R-1, k, t) \\
\ell_{j+R-1,j,t} &= g_{j,t} \\
\ell_{j+R-1,j+R,t} &= \textstyle\sum_{k=1}^{t} \sum_{m=1}^{j} \sum_{n=j+1}^{R} x_{m,n,k} P(m, j+R-1, k, t)
\end{aligned}$$

for outgoing flows, respectively. By means of these quantities, link travel times and transition probabilities can be estimated. For this purpose, the pedestrian velocity-density relation proposed by Weidmann shown in figure 4 is employed.

As this relation applies to one-way traffic only, it needs to be corrected for bi-directional flows. One way to account for that is to reduce the link capacity and thus the pedestrian velocity depending on the flow composition, i.e., the degree of counter flow (Navin and Wheeler, 1969). The link capacity $c_{m,n}$ can be thought of as the minimum effective lane width of a directed link connecting the neighboring nodes $m \to n$. If $w_{m,n}$ represents the walking length of the same

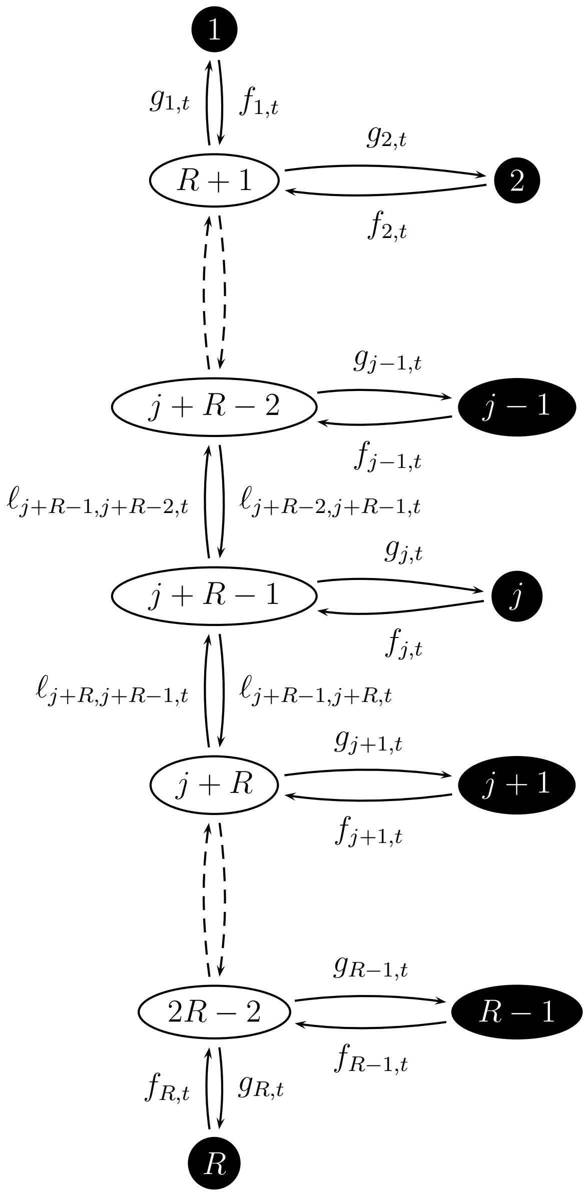

Figure 3: The pedestrian flow network of the train station shown in figure 2, generalized to $(R-2)$ platforms, with R centroids (black) and $(R-2)$ intersection nodes (white).

link, the average pedestrian velocity on link $m \to n$ during time interval t can be expressed as

$$v_{m,n,t} = v\left(c_{m,n}, \ell_{m,n,t}, \ell_{n,m,t}, \tau_{m,n}\right)$$

where the function $v(\cdot)$ implicitly takes into account the velocity-density relation and the capacity reduction due to two-way traffic as described above. $\tau_{m,n,t}$ is a random variable representing arbitrary fluctuations in average walking speed. For the pedestrian network of interest, the trip

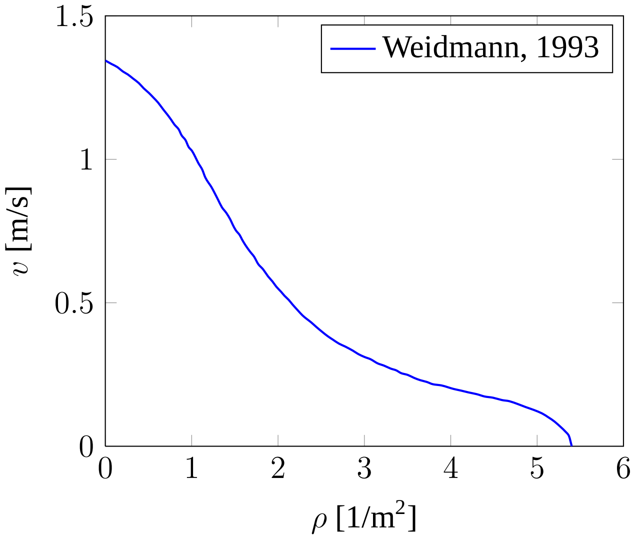

Figure 4: Pedestrian velocity-density relation for planar, unidirectional flow (Weidmann, 1993).

duration from node i to j along path $L_{i,j}$ can then be written as

$$y_{i,j,t} = \sum_{(m,n)\in L_{i,j}} \frac{w_{m,n}}{v_{m,n,t-1+y_{i,m,t}}} \tag{11}$$

For instance, the (unique) path from centroid i to centroid j ($j > i$, w.l.o.g.) can be expressed as

$$L_{i,j} = \{(i, i+R-1), (i+R-1, i+R), \dots, (j-R-2, j-R-1), (j-R-1, j)\}$$

By assessing the distribution of $y_{i,j,t}$, which is a random variable, the transition probability $P(i,j,k,t)$ can be derived. With this, the system of equations is complete and the origin-destination demand of this case study can be approximately computed.

5 Discussion and Outlook

A preliminary methodology has been presented which can be used to predict pedestrian origin-destination demand within a train station when either pedestrian count data or information on the train timetable, track assignment as well as the expected passenger turnover for each train is available. This framework will be refined, extended and implemented for a concrete case study which is similar to the fictitious one mentioned in this work. Beyond this, it will be interesting to generalize the methodology to more complex networks, and to improve the OD estimation procedure by taking pedestrian demand into account which is not caused by trains, but by other transportation modes or intermediate activities such as shopping, eating and so on. In the long term, it might be worthwhile to incorporate such an OD estimation method in a pedestrian dynamics simulator. This would for instance allow to run optimization simulations regarding the effect of train track assignment and train time table on congestion as mentioned in the introduction of this report.

Acknowledgement: This research is supported by the Swiss National Science Foundation grant 200021-141099 'Pedestrian dynamics: flows and behavior'.

References

Cascetta, E. and S. Nguyen (1988) A unified framework for estimating or updating origin/destination matrices from traffic counts, *Transportation Research Part B: Methodological*, **22** (6) 437–455.

Daamen, W. (1999) Modelling pedestrians in transfer stations, paper presented at *Five Years Crossroads of Theory and Practice. Proceedings 5th TRAIL Annual Congress 1999 – Part 2*.

Daamen, W. (2002) SimPed: A pedestrian simulation tool for large pedestrian areas, paper presented at *Conference Proceedings EuroSIW*.

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

Hoogendoorn, S., P. Bovy and W. Daamen (2002) Microscopic pedestrian wayfinding and dynamics modelling, *Pedestrian and Evacuation Dynamics*, 123–155.

Navin, F. and R. Wheeler (1969) Pedestrian flow characteristics, *Traffic Engineering*, **39**.

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