D2.2: Multi-layer passenger flow network model

Project acronym:	TRANS-FORM
Project title:	Smart transfers through unravelling urban form and travel flow dynamics
Funding Scheme:	ERA-NET call on Smart Cities and Communities (ENSCC)

Authors

Riccardo Scarinci	riccardo.scarinci@epfl.ch
Oded Cats	o.cats@tudelft.nl
Johanna Törnquist Krasemann	johanna.tornquist.krasemann@bth.se
Flurin Hänseler	f.s.hanseler@tudelft.nl
Nicholas Molyneaux	nicholas.molyneaux@epfl.ch

Internal Reviewer

Marco Laumanns	mlm@zurich.ibm.com

State:	Final version
Distribution:	Confidential

Deliverable History

Date	Author	Changes
16-11-2016	Johanna Törnquist Krasemann	Draft table of contents
23-11-2016	Johanna Törnquist Krasemann	Updated draft and inserted deadlines. Draft of section 1, 2.1 and 3.1.
7-12-2016	Flurin Hänseler	Added first draft of sections 2.3 and 2.4 (new; to be discussed)
6-1-2017	Oded Cats	Add description of planned model extensions in 3.1
24-1-2017	Johanna Törnquist Krasemann	Update of section for the regional level (2.1 and 3.1.) Updated deadlines and formatting.
20-03-2017	Riccardo Scarinci	Update Section 2.2 and 3
08-05-2017	Riccardo Scarinci	Update outline and time schedule
17-05-2017	Flurin Hänseler	Update of section 2.2, adding of draft of subsection 3.1.2 and section 3.2
22-05-2017	Riccardo Scarinci	Review added material
29-05-2017	Riccardo Scarinci	Introduction review
08-06-2017	Johanna Törnquist Krasemann	Completing the text about the regional level (now section 2.3 and 3.3). Completing section 5.3
09-06-2017	Riccardo Scarinci	Section 3 introduction and review
28-06-2017	Nicholas Molyneaux	Completed Sections 3.1 and 3.4
05-07-2017	Riccardo Scarinci	Add Summary and Conclusion.
07-07-2017	Riccardo Scarinci, Nicholas Molyneaux	Global review
27-07-2017	Marco Laumanns	Internal review
28-08-2017	Riccardo Scarinci	Modification incorporation
29-08-2017	Johana Törnquist Krasemann	Adjusting section 2.3 in line with the review comments.

Contents

Summary

Modelling transportation systems is a crucial aspect that allows the understanding of the internal passenger dynamics and the development of control strategies to improve the public transport service. This deliverable reports the modelling approaches for passenger flows and the related public transport operations used in TRANS-FORM.

We model three distinctive levels: hub, urban and regional. The hub-level model represents the dynamics, movement and activities of passengers in transportation hubs. We present a framework suitable to simulate, evaluate and generate pedestrian management strategies. BusMezzo, a multi-agent and multi-modal simulation, is used to represent the dynamics of public transport systems at urban level. The model includes transit services, travellers and management policies. The regional level is composed of the railway infrastructure and train traffic model. This includes origins and destinations of passengers, lines and stations, and train movement. The train traffic is simulated as a discrete-event model. The model is able to evaluate re-scheduling decisions on arrival and departures times for all trains.

Each level uses information from the other levels. The pedestrian travel times within a transportation hub, generated by the hub level, are used by the urban and regional levels to evaluate passenger dynamics more accurately. Similarly, the hub model uses information from the other two levels to generate precise passenger arrivals. The connection between the regional and urban levels takes place through the hub level. This means that hubs are the interfaces between levels. As a result, this modelling level is placed at the middle of the integration framework.

The models summarised in this deliverable are the foundation for the generation of multi-level control strategies for improving public transport services.

1. Introduction

Modelling passenger dynamics is fundamental to design and manage efficient public transport systems. Work package 2 "Measuring and modelling passenger interchange activities" has exactly this aim.

In Task 2.2 "Linking hubs and urban networks" and Task 2.3 "Moving between networks", at first, we develop the models used to represent public transport system dynamics on a specific level (i.e. hub, urban and regional level). Then, we specify the multi-level interactions among the input and output of these models.

The specific focus of each model is the following.

- Hub level: The interaction between passengers and their movements within a hub.
- Urban level: The interaction between passengers and vehicles within an urban public transport network operated by one or multiple transport service providers including the passenger flow dynamics.
- Regional level: The interaction between primarily passenger trains, but also regional buses, operated by various service providers in regional/national railway traffic system.

These models are then enhanced and developed further by integrating aspects from the other levels in order to better capture the overall system dynamics, particularly related to connecting services and passengers transfers and journeys that cross the different levels.

In this process, we re-arranged the order of the models. Originally, we ordered the models by geographical scale, from regional to urban to hub. However, the interactions among the levels, when described from the passengers' point of view, occur in a different order. The hub level is the intermediate level used by passengers when moving between networks. As a consequence, we use the hub level as an interface between the urban and regional levels.

This deliverable presents the results from Task 2.2 and Task 2.3. In Section 2, we present the modelling approach including how passenger flows are represented on each level independently. In Section 3, we discuss the integration among the three levels. We describe how we extend the models integrating aspects from the other levels, and the purpose and intended use of these. We conclude the deliverable with a discussion about these integrated models and their application in the forthcoming project work in Section 4.

2. Modelling public transportation and passenger flows at each level

The models representing the passenger dynamics at each level are described in the following sections.

2.1. Hub level

In the context of TRANS-FORM, the concept of "Hub" is used in a broad meaning. A hub is defined as any point of interchange between public transport services. Thus it is not primarily related to the size of a station, but rather to its function of interconnecting different services and network layers.

A hub ranges from an entire multifunctional transportation hub, such as a train station with interchange possibilities between trains, metros, trams and busses, to a simple bus stop located on the sidewalk used by one bus line.
Public transport stops (bus/metro/tram/train) used only by one line are considered hubs, because, although no interchanges between two lines can happen at this location, a passenger can transfer between lines walking to different hubs. For example, public transport stops served by a single line located in proximity of train stations. These stops can be used to access larger hubs by walking. This classification is needed to allow the integration with the regional and urban levels.

The global model structure of a hub is represented in Figure 1. In the following, we describe the main components.

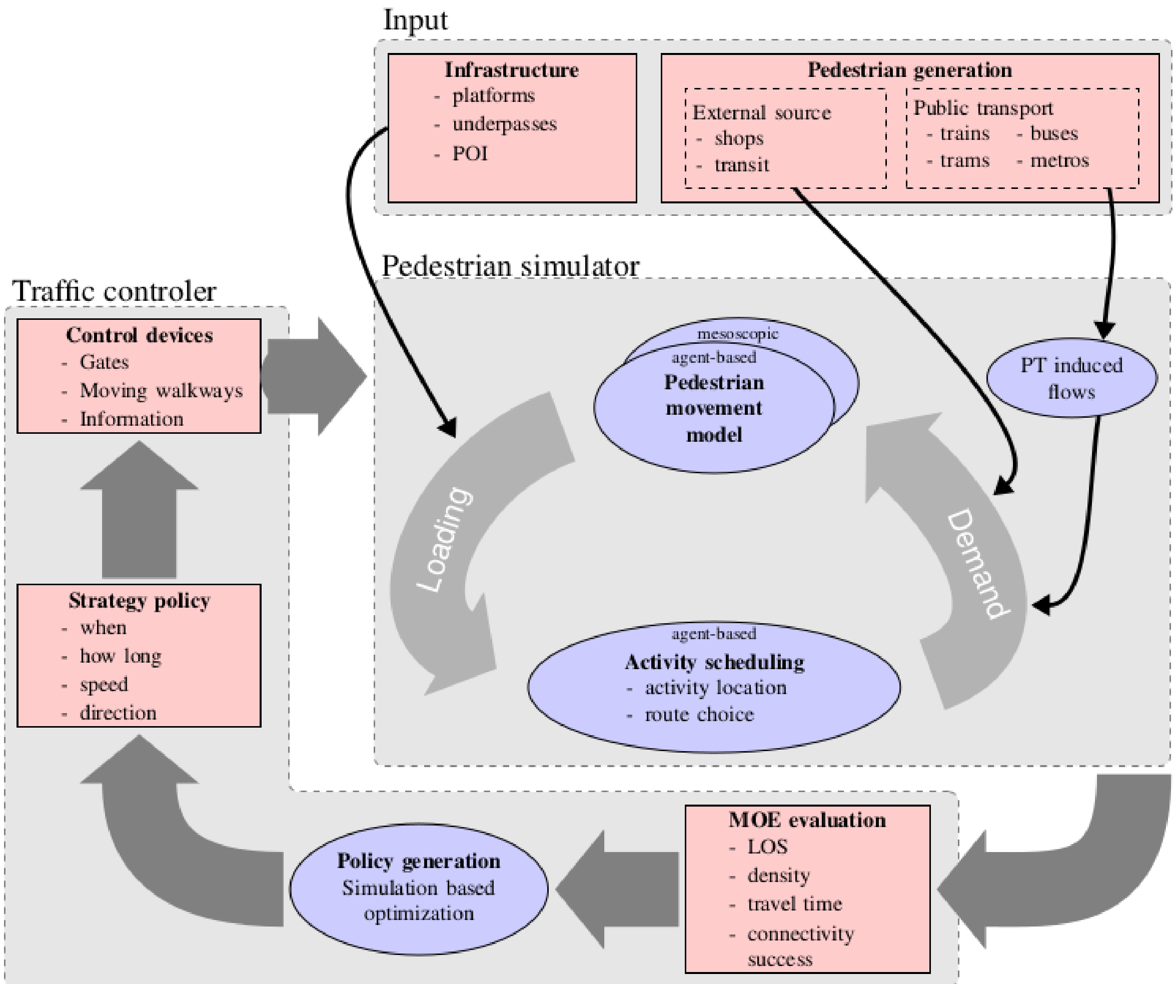

Figure 1 Hub model structure

2.1.1. Inputs

Given the broad definition of hubs, the inputs should be able to represent different types of hubs. The needed inputs for the hub level modelling should cover:

- Infrastructure. The infrastructure could be represented either as a graph or as a geometrical layout that should be processed. We start with a graph representation composed by nodes and arcs. All arcs are directed arcs.

- Demand, including public transport services. The demand can be disaggregate, i.e. a single individual, or aggregated, i.e. a homogenous group of individuals. In both cases, the input information associated with the demand are identical. The only exception is that, in case of an aggregated group, the dimension of the group should be given (e.g. 12 people). The demand needs to have information both at the hub level, and at the network level. Hub level information refers to the hub-specific origin and destination. Meanwhile, Case study level information refers to the entire journey of the individual, independently from how many hubs the individual has used for interchange. The demand should contain information also of hub infrastructure use by individuals that are not necessarily passenger, i.e. individuals that uses the train station for activities not associated to travel. For example, shopping, eating or simply crossing the station to reach other parts of the city. The public transport service is represented by the lines and time table (time or frequency) for each entrance/exit. The representation of this input can be as close as possible to existing standards as, for example, GTFS.
- Management strategies. The set of possible control actions to manage the flow of pedestrian within the hub. These include hard control, such as gates, and soft control, such as information to the passengers.

2.1.2. Modelling

The main internal components of the hub model are the following:

1. Activity choice model and location choice model
2. Movement model
3. Destination choice model
4. Pedestrian flow management strategies
    - Control, information

Activity choice and location choice model

A passenger is generated in the hub with the information defined in the previous section. The activities performed by the passenger are described by the activity choice model. The location where these activities are performed is described by the location choice model. The activity choice is defined based on the required actions a pedestrian must perform (buy a ticket for example) and the optional actions (buy a coffee or newspaper, wait in a restaurant). Once the list of activities is defined, then the activity sequence and location choice is estimated simultaneously. This is required as pedestrians can be considered as utility maximizes (or cost minimizers) and both the location and sequence impact the cost of performing an activity in a specific place.

Movement model

The movement of the individuals and the interactions among individuals is described by the movement model. A route choice model and "traffic" assignment could be necessary depending on the complexity of the hub.
The dynamics within the hub are described in details.
The dynamics among hubs, i.e. the passengers waking from one hub to another hub using the urban network as a connection, are represented with simple assumptions. For example, the travel time of a passenger walking between two hubs is distributed following the flee flow speed distribution plus a delay proportional to the number of intersections between the two hubs.

Destination choice model

A passenger is generated in the hub with an associated (preferred) destination within the hub. This destination is associated to another public transport service in case of transferring passengers. However, the internal hub dynamics could modify this (preferred) destination of the individual.
Example: a passenger goes from Montreux to Geneva by train. The regional-level model says that she takes the train at Montreux at 8:00. She arrives a Lausanne at 8:30. She has the next train at 8:35. She arrives at Geneva at 9:05. She needs to be in Geneva at 9:10 (do we have this information?).

We are interested in the hub dynamics, and if hub level dynamics can influence the (pre-made) regional-level decisions. For example, due to congestion in the underpass, she may decide to take the next train, leaving at 8:45 from Lausanne and arriving at 9:15 in Geneva.
The effects of the hub dynamics to the destination (both at hub level and case study level destination) are evaluated by the destination choice model.

Management strategies

The effects of the management strategies on pedestrian flow and the selection of the optimal control strategy is performed in this module.
We investigate three types of management strategies:

- Control. Control strategies provide physical restrictions of the flows. Example: gates, both hard, e.g. a physical barrier, or soft, e.g. traffic lights for pedestrian.
- Information. Information strategies provide information to the travellers. Examples of information are: information on delay, congestion (both pedestrian facilities and inside the vehicles), not-showing the platform allocation, suggestion of alternative routes. Possibility to evaluate "mandatory" information, e.g. mandatory change of route. This can be enforced in combination with "smart-tickets", i.e. electronic tickets that can be modified by the operator dynamically.
- (soft) Planning. Examples are: layout disposition such as location of tickets machines, information panels, fix flow directions (one-way), e.g. arrows on the floor, etc.

2.1.3. Output

The output of the model is a representation of the dynamics between hubs and inside the hub.

Between hubs

We do not consider congestion for the calculation of the walking travel time between hubs. We assume that Hubs are connected by streets with infinite capacity. We may modify the distribution of travel time (making it wider), or use known information on the connecting streets, such as density/flow, number of intersections and sidewalk width.

- Travel time distribution among all hubs within the walking threshold
- Probability of choosing a different Hub and Service.

Within hubs

Same as between hubs, and in addition

- Travel time distribution between all pairs of entrances/exits (and correlated Level of Service. The Level of Service is related to the density of pedestrian, and it– useful to calculate the total travel satisfaction)
- Probability of choosing a different exit service than the (preferred) destination within the hub when the passenger was generated.

Due to dynamics internal to the hub, passengers could decide to change the destination public transport service. For example, a passenger that arrives at Lausanne train station at 12:28, due to unexpected congestion may lose the connection to his preferred destination Public Transport Service (e.g. the train to Geneva at 12:34). Therefore, the destination of this passenger could be modified. The Hub level gives as an output the probability to choose another destination (exit and Public Transport Service) instead of the preferred one.

2.1.4. Hub classification

A classification of the hub types is needed to distinguish among them. Possible classification criteria are:

- Mode: the number of modes of transport serving the hub:
    - 1-modal, 2-modal, 3-modal, etc…

- Network importance: [0.0-1.0] proportion of passenger using the hub for interchange/access/exit within the network (in the time horizon).
- Line importance: [0.0-1.0] proportion of passenger using the hub for interchange/access/exit within the line (in the time horizon)

2.2. Urban level

The urban level is concerned with modelling the movement of urban transit services, typically referred to as the modelling of urban transit operations, and with describing how travellers make use of them to get from their origin to their intended destination. The latter is known as transit assignment.

In the scope of the present project, an existing urban level model, BusMezzo, is extended to model transit assignment and transit operations by directly taking advantage of the regional model and the hub model. In the following, a brief description of BusMezzo is provided; for an in-depth review, the reader is referred to Toledo et al. (2010) and Cats et al. (2016). For a description of the integration with the aforementioned external model components, see Section 3.

BusMezzo is an agent-based transit assignment and urban transit operations model, considering individual passengers and individual transit services, such as busses or urban trains. The progress of individual passengers is modelled as a sequence of travel decisions which are formulated within a discrete choice framework. An initial choice set generation method provides a set of attractive path alternatives for each origin-destination pair, such that the available choice options are known a priori. When making a decision, passengers take into account the anticipated travel attributes associated with each travel action based on their preferences and expectations of future travel conditions, such as the congestion level of a connecting bus.

Urban passengers are generated randomly based on time-dependent origin-destination demand matrices. The time interval between generations follows a negative exponential distribution. This is suitable in the context of high-frequency services, where passengers do not consult timetables prior to their departures. The generation of passengers at an origin, such as a metro stop, then follows a Poisson process. Passengers stemming from interregional services and their non-uniform generation at system boundaries are considered explicitly by means of "gates." This represents one of the extensions of BusMezzo within TRANS-FORM, and is described in Section 3.

Walking links in the urban level model are not represented explicitly. Instead, for each pair of nodes within the walking network (such as between a bus stop and a train platform), the expected walking time is estimated from their distance. For pairs of walking nodes within a transportation hub, walking time estimates provided by the hub model can be integrated. This is particularly useful for modelling transfer passengers, such as those alighting from a train on "Platform 7" headed to "Bus Stop B." In this way, realistic transfer times are ensured, taking into account the dynamics of way finding or congestion within stations. The possibility to use exogenous walking time estimates within urban transit assignment represents another dedicated extension of BusMezzo within Trans-Form.

The progress of public transport vehicles between stops is modelled with a mesoscopic traffic simulation model, allowing to model the operation dynamics of large-scale transit systems. Dwelling times at stops are determined based on the interactions between transit services and passengers. Different public transport modes, such as metro, light rail, commuter trains and buses, have distinct vehicle types, operating speeds, travel time variability, and may be operated with different holding control strategies. Each transit service is assigned with a chain of trips, allowing to capture the dependency between successive trips and the potential propagation of delays from trip to trip.

Riding times between stops are composed of running times on links and delays at intersections, which are computed based on speed-density functions and stochastic queueing models. For that purpose, road links are divided into a running part, containing vehicles that are not delayed by downstream capacity limits, and a queueing part, extending upstream from the end of the link when capacity is exceeded. The boundary between the running and queueing parts depends dynamically on the length of the queue. Alternatively, running times may be inferred from a priori known distributions, obtained for instance from automatic vehicle location data.

The resulting model allows to describe the dynamics of passengers and vehicles within a public transport network, to assess its reliability and performance, and to evaluate management policies for instance in case of disruptions. In particular for the latter, the integration with the hub and regional models are key, and are further described in Section 3.

2.3. Regional Level

Public transportation on a regional level encompasses - in this project - regional and interregional public transport services including the interface between regional, subsidized services and national commercial train services. The backbone of the regional public service network is in this project assumed to be a railway traffic network and the scheduled passenger train services. A network of selected bus services are then added to cater as complementary services and to potentially also connect the train service network with the urban level (see Section 2.2) via hubs (see Section 2.1).

We originate from an existing mesoscopic optimization model of a railway traffic system with the intention to extend and develop it further. This model was initially developed for railway traffic re-scheduling and delay management, and it is described in detail in (Törnquist and Persson, 2007), (Törnquist, 2007) and Törnquist Krasemann (2015). This existing optimization-based re-scheduling model is composed of three main parts:

- An infrastructure model including the configuration of the lines, the stations, capacity restrictions and interaction scheme.
- A traffic model including the timetables and train service properties (including freight trains).
- A passenger flow model, where passenger flow data is aggregated into a set of different passenger groups each associated with specific train service departures, arrivals and pre-defined train connections.

These parts and the intended extensions are explained below.

2.3.1. Infrastructure and train traffic model

The railway network is modelled with the detail level according to Figure 2.3.1. below.

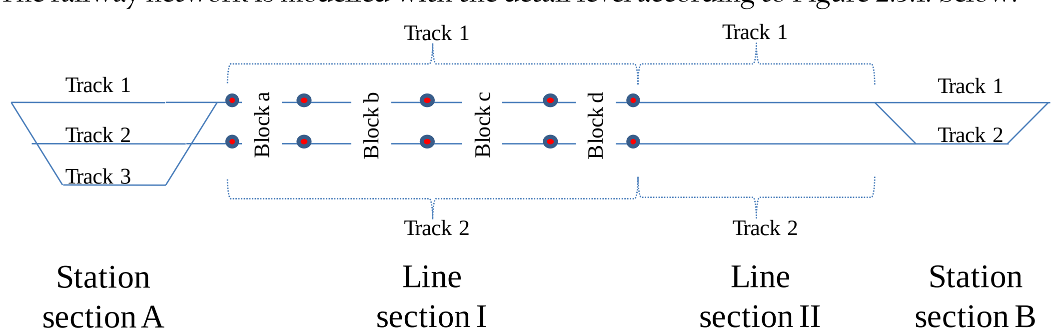

Figure 2.3.1. Railway network representation.

The network is modelled mainly as two types of resources: Line sections and station sections and they are modelled slightly different. Each section has a finite capacity (i.e. a number of available, individual tracks) and each unit of capacity can hold at most one train at a time. In addition there are constraints to separate the trains sufficiently in time to respect safety and braking distances. If a line section corresponds to a junction connecting e.g. two station tracks into a single-tracked line section, it has no real capacity, but still needs to be modelled since it is a critical scarce resource for transferring trains from one track to another. The capacity is in those special cases set to '0' and the trains are using a "virtual track" for their instant passage and cannot stop there for obvious reasons. For station sections, each individual track has certain properties including a specific track length and platform length.

The model of the train traffic is a discrete-event model. Each train *i* has a set of train events, following the specification of the original timetable. This set is sorted based on time and is referred to as the train event list of train *i*. Each train event corresponds to a train taking the possession of a track resource on a specific section. If the section has multiple tracks, there is a preferred track of each event, but any of the available tracks can be allocated and used by that event as long as the constraints are satisfied w.r.t. track and platform length.

The re-scheduling model includes decision making regarding real-time allocation of arrival and departures times for all trains when deviations from the timetable occur as well as track and platform assignment, and train prioritization when resource conflicts arise.

In the model of the train traffic we also want to model and consider relevant dependencies between certain trains at certain stations. These dependencies can be "strict", which is the case when there are rolling-stock circulations forcing one train service to "hold" at its origin destination until the associated physical train has arrived to that specific station. These can also be "soft", which is the case when there are connecting trains for scheduled passenger transfers, i.e. *connections*.

In the model, there are binary variables used to capture if each scheduled connection is successful, or not, depending on the re-scheduling of the train arrival and departure times. That is, if the time between the arrival of the incoming train and the departure of the outgoing train is large enough to enable a smooth passenger transfer. The minimum required time for the associated passenger transfer is dynamic since it is dependent on the allocated platforms for the connecting trains.

As the existing model focuses on real-time train traffic management, it does not include management of regional bus transport services that are relevant for the passengers to consider during disturbances. This extension is briefly described below and will be described in more detail in the forthcoming deliverable "D3.1: A toolbox of real-time strategies for smart transfers".

2.3.2. Passenger flow and transfer model

In the existing optimization-based re-scheduling model, there is only very limited focus on passenger delays and transfer dynamics. The passenger flow was previously mainly represented as a static O-D matrix (i.e. assigned OD demand) which specifies the number of passengers alighting a specific train at a specific station, see further Törnquist (2007). The impact of delays on passengers thus only occurs and is penalized at the moment those passengers alight the delayed train. In another version of the model, see (Törnquist and Persson, 2005) penalties for breaking connections are also considered, and those penalties are based on the assumption that we know how many passengers that transfer from one train to another at selected transfer stations. If the passengers miss their connection, they all have to wait "x" minutes until the next corresponding train service arrives. This is a valid assumption if there are no relevant, alternative services to select.

However, we now extend the existing re-scheduling model to also include a selection of regional bus services, which may be an alternative public transport service option if certain train connections and passenger transfers will fail. Those bus services are modelled similarly as the train services, apart from that there are no infrastructure capacity limitations considered at the bus stations (hubs), or on the road network connecting the bus stations. Given the alternative train and bus services, we need to model the passenger group-specific alternative routes and the passengers' dynamic preferences.

The passenger flow is now aggregated into and modelled as a set of passenger groups. A group can be seen as a cluster of a fixed number of passengers, whom want to pursue the same OD-trip during the same time period. Each passenger group therefore has a dedicated initial transport service (i.e. a scheduled departure time with a specific public transport service at a specific origin station) and a preferred arrival time at a specific destination with a specific transport service. The initial and final service can be the same and then the preferred route is a direct route. The route may also contain a pre-defined sequence of specific bus/train connections, which then implicitly defines the sequence of transport services used and the expected travel time, waiting time and final arrival time.

Input to the model is a pre-computed limited set of relevant routes for each passenger group, including the preferred route. All routes available for a certain group start at the same location, but the routes may start and end with different transport services depending on time of day and which connections that are included in the route.

During the re-scheduling, a route can only be dynamically selected by the group if all included connections are successful, which is dependent on the dynamically assigned service arrival times, departures times and platforms that the optimization-based re-scheduling model generates.

3. Integrating the levels

In this section, we describe how the three aforementioned model components are extended. We also discuss how the output from a level is used as input from the other levels.

We chose to order the models based on the passenger perspective. For example, a passenger performing an intermodal trip starts the trip at a transportation hub (hub level), embarks a long distance train (reginal level), then she performs an exchange to urban bus (urban level) at another hub (hub level) until she reaches her final destination.

Looking at the order of the levels from a geographical perspective, i.e. the scale of the level, it is natural to order them from the larger to the smaller, see Figure 2 (a). However, if we look at the order from the passenger perspective, the hub level is in the middle, see Figure 2 (b). Hubs are used by passengers as transfer place within the same level and to other levels. For this reason, we structure the integration among the levels placing the hub in the middle.

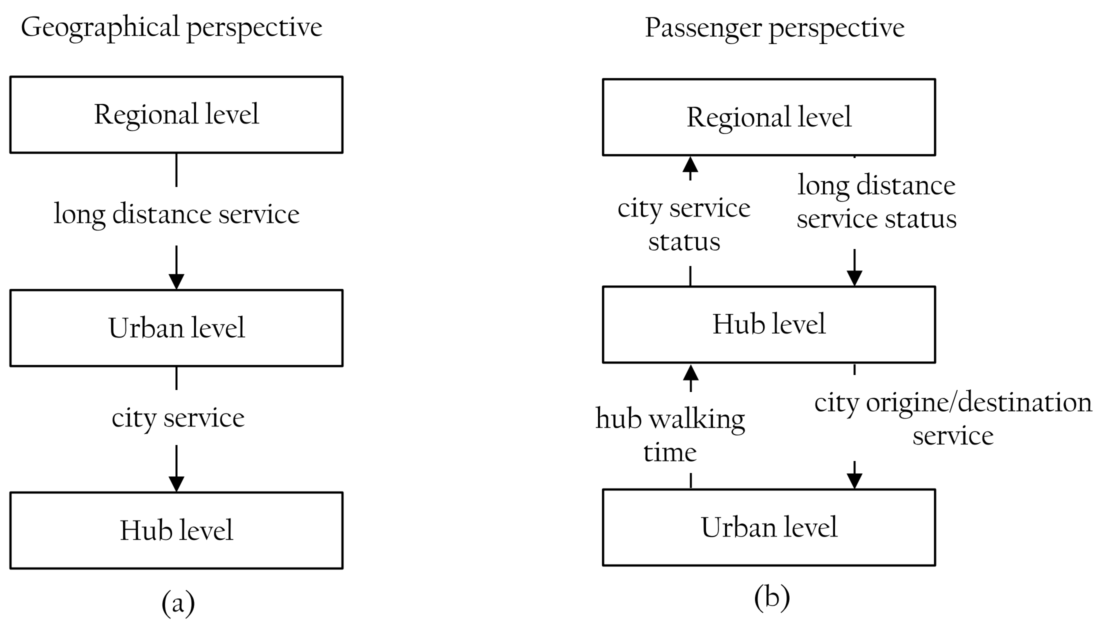

Figure 2 Level order based on the geographical perspective versus the passenger perspective

In the following subsections, we describe two main components for each model: the outputs and the use of the inputs. The outputs of a model are the information generated inside the model that are used by the other levels. These are represented by arrows in Figure 2. From the point of view of the receiving model, these are inputs. The second component that we describe is how each model uses these inputs internally.

First, we describe the outputs of the hub level in Section 3.1. Then, for the urban and regional level, we report both their outputs and how they use the inputs from the hub model in Section 3.2 and Section 3.3, respectively. Finally, Section 3.4 describes how the hub level uses the inputs from the other models. Note that it is necessary to split the description of the hub level into two separate sections in order to present the hub level outputs before describing how these are used by the other levels.

3.1. Hub level to Regional and Urban level (EPFL, 16/06/2017)

The hub model provides expected walking distances and walking time distributions to the urban and regional models, such that potentially time-dependent walking time distributions for each relevant OD relation inside the station are available for off-line use. The justification for the offline estimation is based on the fact that walking times within stations are typically an order of magnitude shorter than entire transit trips, and that their temporal variation, even in case of pedestrian congestion, is relatively small.

Walking time distributions are described by their quantiles. As compared to a parametric description of distributions, this allows the specification of general distributions, and in particular of distributions that are bounded and asymmetric.

Quantiles are defined as follows. The k-th quantile of a set of walking times divides them so that k % of the walking times lie below and (100 - k)% of the walking times lie above. The number of quantiles, and their position, can be chosen freely. By convention it is assumed that quantiles are sorted in increasing order.

Importantly, walking times are not assumed to be symmetric in the sense that walking from A to B may be associated with different walking times than walking from B to A. Likewise, walking time distributions may depend on time, and thus the time interval for their validity has to be specified. The latter is mostly included for flexibility in future extensions, and for most cases of limited importance, unless serious crowding occurs.

Besides walking times, also walking distances are provided. Specifically, for each origin-destination pair, the expected walking distance is provided; if no pedestrian is associated with an OD pair, instead the minimum walking distance is provided. Walking distances are thus provided for all pairs of stops, except those for which walking is considered infeasible. Unlike in the case of walking times, for walking distances only a scalar value is provided.

An example of the JSON specification is available in the appendix of this document, in section 5.1.

3.2. Urban level to/from Hub level

BusMezzo is developed as a stand-alone scientific tool that is used also beyond the scope of Trans-Form. As such, for data input and output a general format is used. Specifically, the specification of walking times, walking distances and transit schedules is integrated in a "transit_network" parameter file that defines also many other properties of the transit network, such as transit stops or transit lines. Given this general-purpose parameter format, the reading of the output of the hub and regional level by the urban level requires some pre-processing, typically in the form of copying and pasting the desired information in the corresponding section of the transit network parameter file. Likewise, the output of the urban model, such as disaggregate demand at the level of hubs, needs to be extracted a posteriori by means of a semi-automatic process.

3.2.1. Output to the hub level

BusMezzo provides detailed passenger information at an individual level, which can be used to infer pedestrian origin-destination demand at the level of transportation hubs. Importantly, for connecting passengers the feeding and connecting transit runs are provided; for outgoing or incoming passengers (for instance arriving or leaving by walking from/to home), besides their transit service also their initial/final "stop" is known. This allows to infer the full pedestrian origin-destination demand at any desired transportation hub within the perimeter of the modeled network. Alternatively, the demand at the hub level may be estimated independently from the output of BusMezzo based on data and/or scenario assumptions.

3.2.2. Use of input from the hub level

Two main inputs are used: walking dynamics within the hub and the schedule of transit services.

Walking dynamics within the hub.

The connectivity of stops by walking is important for computing travel times, as well as for path choice generation and path choice decisions. To describe the connectivity, the distance in space, and if available, the distance in time in terms of walking times provided by the hub level are used.

Whenever available, pedestrian walking times of passengers are directly based on the walking time distributions provided by the hub model. This is valid for both intra-hub walking times and inter-hub walking times. Specifically, walking times are drawn from the distribution specified in Section 3.1.2. by generating a random quantile k, and by using linear interpolation between the two nearest quantiles to estimate the corresponding walking time. If for a given pair of stops and time no corresponding specification of the walking time distribution is available, walking times are estimated based on walking distances by assuming a Gaussian walking speed distribution from the literature.

For route choice generation, and for predicting traveller decisions, the provided estimate of walking distance is used instead for two reasons. First, a scalar value is easier to handle than a quantile-based distribution for choice set generation and travel decisions, which also need to consider downstream consequences. Specifically, besides a lower computational cost, the use of a deterministic value avoids the introduction of undesired fluctuations in the learning process, as walking times typically fluctuate strongly. Second, for "unlikely" walking routes, an estimate of travel time may not be available in the first place, whereas the (minimum) walking distance can always be computed.

Schedule of transit services.

The urban model takes the schedule of transit services as input, allowing for different formats of specifications. One format considers the explicit specification of transit service times, as it is typical e.g. for train services. The transit schedules produced by the regional model for instance are directly incorporated in the aforementioned "transit network" parameter file and treated no differently than any other schedules of urban transit services.

A special case pertains to supra-urban transit services, such as inter-regional trains. A majority of their stops may lie outside the area of interest. Nevertheless, passengers originating or terminating at such stops may influence the dynamics and performance of the considered urban system. To account for that, the notion of gates is introduced. A gate is a node that represents stops associated with a specific transit line that lie outside the area of interest. In gate nodes, passengers are generated based on the timetable a "few moments" before the next service runs, instead of randomly as in other stops. This avoids the introduction of artificial waiting times.

In summary, the urban model makes use of the deterministic train service times estimated by the regional model, and considers passenger demand of long-distance travelers by means of gate nodes, representing unmodelled remote stations. The explicit consideration of train rescheduling will allow amongst others to study transfer synchronization control strategies in case of disruptions, as planned within WP3 of the Trans-Form project.

3.3. Regional level to/from Hub level

Figure 3.3.1 below illustrates the intended interaction between the optimization-based re-scheduling model for regional public transport management, and the hub level. The hub level provides detailed transfer times between pairs of platforms, and it is the main interface between the regional and the urban level since a passenger flow exchange between those levels may occur via common platforms within the hubs. See section 5.3 and deliverable D3.1 for more detailed information about the intended data exchange between the models.

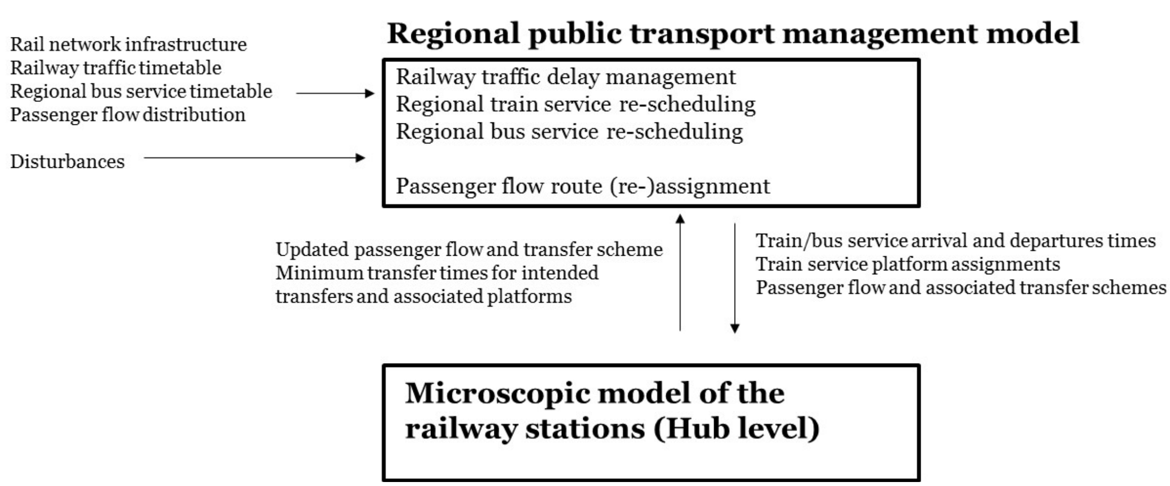

Figure 3.3.1. Overview of the model of regional public transport management and its connection to the hub level.

3.4. Hub level from Regional and Urban level

The hub model uses the output from the urban model and regional model in a similar way: the data provided by the transit models is used as pedestrian demand for the hub model. The walking times of pedestrians inside the hub (being transfer or non-transfer passengers) is strongly dependent on the congestion inside the hub. Furthermore, congestion is highly variable in time and space as public transport services (trains, buses, trams, metro) unload many passengers in a short time span and in a small area of the hub.
The information provided by the other models allows the consideration of accurate pedestrian flows inside the hub as the precise arrival times of a trains (for example) are known. This provides great insight as the highly localised (space and time) congestion can be modelled. Without this information, only average flows for the hubs are known. Moreover, the destination inside the hub is known thanks to the other models as the target public transport service is known.
Finally, when rescheduling takes place, the knowledge of the arrival and departure times of the services allows an estimation of problematic level-of-service through the simulations accomplished by the hub model. If the schedule following a disruption plans many arrivals at the same time, very high congestion can occur and this can lead to critical situations.

3.4.1. Use of input from the Regional level

The regional rescheduling model works with aggregate flows. Passengers aren't modelled individually, but groups of passengers are considered. As the hub model currently uses a disaggregate approach to modelling pedestrians, these flows must be broken down into individual agents with their own characteristics (origin, destination, entry time, etc). This process (train induced flows) then generates individual pedestrians.

3.4.2. Use of input from the Urban level

The urban model provides agent-based information about passenger flows. Each passenger is modelled individually with highly detailed information about his trip. Nevertheless, the alighting process from public transport vehicles needs still to be modelled, hence the train induced flows must still be considered.

4. Conclusions

We present the modelling approaches used to represent passenger and public transport service dynamics. Three different models are developed, one for each level: hub, urban and regional.

The regional model (Section 2.3) is used to compute train arrival/departure times in the hub of interest, subject to a scenario of interest (peak hour, disruption, etc.). This pre-computation is mainly justified by the fact that train operations are rarely adjusted when service irregularities occur at the urban level.

The hub model (Section 2.1) provides pre-computed walking times based on the same scenario, such that time-dependent distributions for each relevant OD relation inside the station are available for off-line use. The justification of that pre-computation is based on the fact that walking times within stations are typically an order of magnitude shorter than entire transit trips, and that their variation, even in case of pedestrian congestion, is relatively small.

The urban model (Section 2.2) is the parent model. When estimating pedestrian walking times in stations, it draws from the pre-computed, OD-specific walking time distributions provided by the hub model whenever needed. Similarly, it makes use of the (deterministic) train arrival/departure times estimated by the regional model.

The three models are not integrated in a single integrated model, but they interact with the use of input/output exchange. In the next phases of TRANS-FORM project, this interaction among the three models will allow an indirect evaluation of the consequence of multi-level control strategies. The Hague, The Netherlands will be used as common case study for this task.

5. Appendix

5.1. Hub level output specification

Below a short sample of the output specification from the hub level. This is a JSON array, where each element contains the following fields:

- "o": String
- "d": String
- "start_timestamp": String
- "end_timestamp": String
- "quantiles": JSON array of integers
- "values": JSON array of doubles
- "sample_size": integer
- "mean_walking_distance": double


```
[ {
"o" : "lausannePl34WestRamps",
"d" : "lausannePl56WestStairs",
"start_timestamp" : "1970-01-01 00:00:00",
"end_timestamp" : "2100-01-01 00:00:00",
"quantiles" : [ 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 ],
"values" : [ 34.2, 34.3, 37.1, 43.8, 45.8, 47.1, 52.6, 54.5, 65.2, 70.0, 77.9 ],
"sample_size" : 11,
"mean_walking_distance": 51.265
}, {
"o" : "lausannePl34WestStairs",
"d" : "lausannePl56WestStairs",
"start_timestamp" : "1970-01-01 00:00:00",
"end_timestamp" : "2100-01-01 00:00:00",
"quantiles" : [ 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 ],
"values" : [ 31.9, 33.4, 34.5, 38.23, 38.4, 40.0, 41.4, 42.6, 43.1, 44.3, 72.51 ],
"sample_size" : 12,
"mean_walking_distance": 41.69
} ]
```


5.2. Urban level output specifications

A complete specification of all inputs and outputs of the urban model are available in a separate document, entitled BusMezzoInput/Output Formats, that can be found on DropBox:

https://www.dropbox.com/s/gcy4wpe8wg0tzg5/BusMezzo%20Input-Output%20format%20v2.0.docx?dl=0

5.3. Regional level output specifications

The output contains a revised timetable with a similar structure of data as the initial train traffic timetable. It contains updated information about all events (i.e. transport service activities) in the data instance. The following data per line is included, where the last three items contains the updated data:

- Event type {TSP = Train stop, TMV = Train movement}
- Section index {0,1,...,n}

- Train/bus number
- Event number (in the complete event list of that train/bus, not only in this data instance), {1,2,..., m}
- Section name
- Announced start time (double), counted in seconds from midnight that date.
- Announced end time (double), counted in seconds from midnight that date.
- Minimum running time given in seconds (for stations it corresponds to the minimum dwell time, which is > 0 for commercial stops).
- Planned track/platform (track index)
- Number of passengers alighting (integer value)
- New start time (double), counted in seconds from midnight that date.
- New end time (double), counted in seconds from midnight that date.
- Allocated track/platform (track index)

The output data concerning the predicted passenger flow is specified by the following lines starting with the string "PAX" followed by:

- Train/bus number
- Section name
- Section index {0,1,...,n}
- Total number of passengers alighting the train/bus at that station

PAX 1046 CK 0 20

If one also wants to have passenger transfer data concerning specific connections, the auxiliary data field in the "TAS data" below can be used to give the number of transferring passengers.

Pre-defined dependencies between certain train and bus services at certain stations are referred to as "Train Associations" – TAS. These can be "strict", which is the case when there are rolling-stock circulations forcing one train service to "hold" at its origin destination until the associated physical train has arrived to that specific station. These can also be "soft", which is the case when there are connecting trains for scheduled passenger transfers. The data includes information about all pre-defined *directed* dependencies, one per line starting with the string "TAS":

- Section name (the station section, where they "connect")
- Section index {0,1,...,n}
- Train number of the feeding, "main" train
- Arrival time of the main train at the specific station (in format hh:mm:ss)
- Train number of the connecting train
- Departure time of the main train from the specific station (in format hh:mm:ss)
- Type of connection {0=soft; 1=strict}
- Auxiliary data (String)

TAS CK 0 104615:12:00 1091 15:47:00 1 X

If the connection is bi-directional, the train association contains two dependencies – one per line.

6. References

Eliasson, J. (2016). A dynamic stochastic model for evaluating congestion and crowding effects in transit systems, *Transportation Research Part B: Methodological* 89: 43–57.

Toledo, T., Cats, O., Burghout, W. and Koutsopoulos, H. N. (2010). Mesoscopic simulation for transit operations, *Transportation Research Part C: Emerging Technologies* 18(6): 896–908.

Törnquist, J., Persson, J.A., (2005) Train traffic deviation handling using Tabu Search and Simulated Annealing, *in Proceedings of HICSS38, Big Island, Hawaii, January 2005.*

Törnquist, J., Persson J., (2007), N-tracked railway traffic re-scheduling during disturbances, *Transportation Research Part B: Methodological 41 (3):* 342-362.

Törnquist, J. (2007). Railway traffic disturbance management-An experimental analysis of disturbance complexity, management objectives and limitations in planning horizon. *Transportation Research Part A: Policy and Practice No. 41 (3), 249-266.*

Törnquist Krasemann, J. (2015) "Computational decision-support for railway traffic management and associated configuration challenges: An experimental study", *Journal of Rail Transport Planning & Management*, Elsevier, Volume 5, Issue 3, November 2015, pp.95–109.