SIMULATION OF CONGESTION PHENOMENA AND STRATEGIC PASSENGER BEHAVIOUR ON TRANSIT NETWORKS

Lory Michelle Bresciani Miristice[1], Dario Menichetti[2], Guido Gentile[3]

[1] *Università di Roma La Sapienza*
Via Eudossiana 18 - 00184 Roma, Italy
+39 335 1502 532, lory.bresciani@ptvgroup.com

[2] *Università IUAV di Venezia*
Ca' Tron Santa Croce 195 - 30135 Venezia, Italy
+971 50 6023 320, dario.menichetti.ing@hotmail.it

[3] *Università di Roma La Sapienza*
Via Eudossiana 18 - 00184 Roma, Italy
+39 06 4585 150, guido.gentile@uniroma1.it

This article focuses on the simulation of transit networks, including passengers' congestion phenomena and strategic behaviour, as a tool for planning and operation. Specifically, it describes a model of user equilibrium on large-scale transit networks considering on-board overcrowding, queuing at stops and dwelling delay, and representing in addition the effects of random events, such as the availability of seats and the waiting of vehicles, through hyperarcs. The assignment algorithm computes a strategy-based user equilibrium with implicit hyperpath enumeration by solving a fixed-point formulation, through a gradient projection method, which highly improves convergence performance. The model was implemented in .net as a dll that uses the software PTV Visum as I/O source. This article will focus on the model formulation and its capability of reproducing the target phenomena, while the results of model validation and testing will be described in a following paper.

Keywords: public transport assignment, hyperpaths, user equilibrium, failing probabilities, common lines

1. Introduction

1.1. Motivation

According to the World Health Organization, on the one hand, transport has a positive key role in the economy, on the other hand, its externalities can harm human health and the environment. In the European region alone, about 100,000 premature adult deaths occur each year due to air pollution, which is mostly caused by emissions from road traffic. In addition, road accidents result in about 127,000 deaths and 2.4 million injuries per year, killing more young people aged 5-29 than any other cause. Moreover, transport is the fastest growing source of fossil-fuel CO_2 emissions, which is one of the main factors of climate change. Finally, traffic noise and congestion damage health, psychological well-being, work performances and overall life satisfaction (Dora and Phillips, 2000).

Public transport is way more efficient of private transport in terms of the above externalities as well as in terms of space usage: just consider that one metro line can carry more than 50,000 passengers per hour, while one road lane can serve at most 2,000 vehicles per hour. Moreover, in the near future ride sharing with electric and autonomous vehicles will probably change our mobility habits but will not solve the capacity issue. Therefore, governments are still recommended to invest on public transport, attempting to attract more users by increasing the quality of transit service in terms of comfort, reliability and speed.

However, also public transport may be negatively influenced by congestion, especially in presence of the desired higher modal shift. Oversaturation of vehicles makes boarding/alighting at stops a difficult operation, causing significant delays and heavily affecting service regularity, while the stress of riding on overcrowded transit carriers may reduce people's productivity at work. For instance, in London transit network, congestion costs to passengers around £230 million per year (Oxford Economic Forecasting, 2003). Level of service increase can be achieved by reducing congestion through careful planning of the public transport service. To this purpose, different transit assignment models have been introduced in the last 50 years.

This article focuses on the simulation of public transport networks, including passengers' congestion due to vehicle and platform capacities, and strategic behaviour due to en-trip unpredictable events, such as failing to sit or boarding and vehicles arrivals at stops. In particular, the following typical transit phenomena will be modelled:

- passengers discomfort due to overcrowding on-board of a vehicle and at platforms;
- passengers queuing/waiting at stops for more than one vehicle arrivals, due to the inability of boarding crowded carriers;
- dwelling delay for passengers boarding and alighting, due to limited door capacity and vehicle overcrowding;
- limited availability of seats, for boarding and dwelling passengers;
- waiting for the first vehicle of an attractive set of lines at stops.

1.2. Existing approaches to transit assignment

The available approaches to the simulation of public transport networks and transit services can be divided into three main categories: frequency-based (FB) models, schedule-based (SB) models and agent-based models (AB).

SB models are usually adopted to simulate low-frequency and high reliability services, which have timetables available to users with departures and arrivals times at each stop. The typical implementation of this approach requires to introduce a diachronic graph to represents every single run of the service, making the models inherently dynamic. Indeed, the assignment results (travel times and passenger flows) are always referred to a specific run and thus to a certain time of the day. SB models can represent in details passengers' congestion due to vehicle capacities but are not suitable to simulate the delays of vehicles due to alighting and boarding passengers at stops, as the diachronic graph is conceived as a fixed space-time network. Another weakness of the diachronic graph is the representation of the time-continuous elements of the transit network, such as the pedestrian links and the interfaces with multimodal transport systems, which may result in a relevant growth of the problem dimension in terms of arc numbers.

FB models are usually preferred to simulate high-frequency or low-reliability services, where vehicles are not able to follow timetables or when timetables are not considered by users. These models simulate services as they are perceived by passengers, that is in terms of probabilistic departure events. All runs of the same transit line are considered by passengers as a unitary supply facility with a certain frequency, thus users do not try to synchronize their arrival at stops with the scheduled passage of carries (if any). In this case, it is rather convenient to wait for a certain set of attractive lines and board the first arriving vehicle of them, because this allows to reduce significantly the waiting times, that are a major component of urban journey cost. FB models rely on a spatial/functional graph without usually involving dynamic aspects. Within the limits of static assignment, it is possible to simulate (as it will be shown in this paper) most of the relevant phenomena occurring on transit networks, including non-separable congestion and strategic behaviour.

AB models for transit assignment adopt essentially a meso or micro simulation approach, with its pros (e.g. great detail with individual vehicles and passengers) and cons (e.g. high runtime, difficult calibration and convergence issues). Because of these problems FB models are still the most used in practice.

The state-of-the-art of transit assignment is summarized in the recent book by Gentile and Noekel (2016), which provides a complete review of the available approaches to the simulation of public transport networks, as well as the available methodologies to encompass emerging aspects such as fleet control, passenger information and multimodality.

The FB model presented in this article derives from a combination of the transit assignment models introduced in Gentile *et al.* (2016a) and Gentile *et al.* (2016b), providing a practical implementation of the equilibrium model and a detailed discussion of its supply side.

2. User equilibrium and network loading

2.1. Fixed-point formulation

The transport network is represented as a directed graph (N, A), where N is the set of nodes and A is the set of arcs, each one with its attributes. Travel demand is segmented into a set G of user classes, each one with its preferences. The study area is partitioned into a set of traffic zones $Z \subseteq N$, each one associated with a specific centroid node, where trips start and end.

A user equilibrium for the transit assignment model discussed in this paper can be found by solving the fixed-point problem with implicit strategy enumeration schematized in Figure 2.1, whose main variables are:

q_a^{dg} flow on arc $a \in A$ of class $g \in G$ users directed to destination $d \in Z$ - it is the main variable;

q_a volume of arc $a \in A$;

c_{ag} cost of travelling on arc $a \in A$ for class $g \in G$ users;

π_{ag} diversion probability of arc $a \in A$ for class $g \in G$ users;

p_a^{dg} conditional probability of arc $a \in A$ for class $g \in G$ users directed to destination $d \in Z$;

w_i^{dg} expected cost of class $g \in G$ users from node $i \in N$ to destination $d \in Z$;

d_{odg} demand trip flow of class $g \in G$ users from origin $o \in Z$ to destination $d \in Z$;

y_a^{dg} auxiliary flow on arc $a \in A$ of class $g \in G$ users directed to destination $d \in Z$.
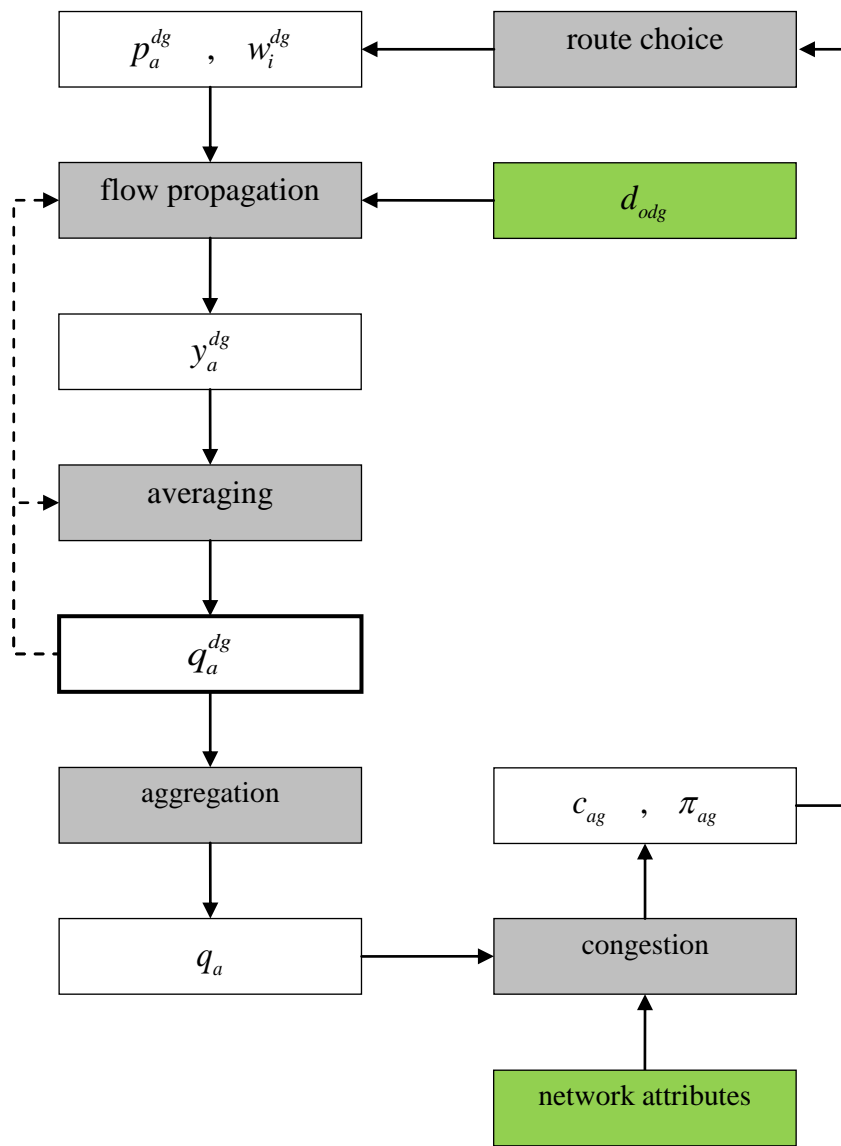

Figure 2.1. Fixed-point schema of the equilibrium model

The main peculiarity of this model is that congestion affects not only the arc costs but also the diversion probabilities, that are the results of random events on the network.

While the network loading model (route choice and flow propagation) is clearly partitioned into demand segments (group and destination), the congestion model is non-separable and requires flows to be aggregated into volumes.

To cast the proposed fixed-point schema into an efficient equilibrium algorithm, arc costs and placing probabilities must be recomputed through the congestion model right after the update of the destination specific flows for each class (without waiting for all the other demand segments). Moreover, it is convenient to repeat the optimization of the same problem partition until the convergence to equilibrium is sufficiently improved, before analysing the next demand segment.

2.2. Route choice model

The main feature of the proposed strategy-based model for transit assignment is the presence of particular nodes where the conditional probabilities are not outcome of passengers' choices but depend on the average results of network events. Indeed, at stops $S \subseteq N$ passengers board on the first available attractive line, while at placing nodes $P \subseteq N$ they are affected by the probability of seating or boarding. Moreover, these probabilities are affected by congestion. The route choice model has then a different form depending on the node type.

For a standard node $i \in N-(S+P)$ it is:

$$w_i^{dg} = Min\left(c_{ag} + w_{N_a^+}^{dg} : a \in A_i^+ \cap A_d\right), \tag{2.1}$$

$$s_i^{dg} = ArgMin\left(c_{ag} + w_{N_a^+}^{dg} : a \in A_i^+ \cap A_d\right), \tag{2.2}$$

$$p_a^{dg} = B\left(a = s_i^{dg}\right), \quad \forall a \in A_i^+, \tag{2.3}$$

where the Boolean function $B(x)$ yields 1, if x is true, and 0, otherwise.

This is the classical determinist model where users of class $g \in G$ choose the successive arc $s_i^{dg} \in A$ to reach their destination $d \in Z$ form node $i \in N$ as the best local alternative among the forward star:

- $A_i^+ = \{a \in A: N_a^- = i\}$, where $N_a^- \in N$ and $N_a^+ \in N$ denote, respectively, the tail and head nodes of arc a;
- $A_i^- = \{a \in A: N_a^+ = i\}$ will denote instead the backward start of node $i \in N$.

Note that in the proposed model we limit the route choice to the so called *efficient* arcs A_d of each destination $d \in Z$. More specifically, we consider only arcs whose head is closer to the destination than the tail in terms of a given graph topological order $z(i, d)$, which essentially measures a distance from each node $i \in N$ to d:

$$A_d = \{a \in A: z(N_a^-, d) > z(N_a^+, d)\}. \tag{2.4}$$

The restriction to efficient arcs greatly simplifies the whole equilibrium model as it allows to easily compute route choice and flow propagation of each destination by processing nodes in reverse and direct topological order, respectively. In section 5 we will relief this assumption for the sake of generality.

For a stop node $i \in S$ it is:

$$w_i^{dg} = \frac{1}{\displaystyle\sum_{a \in A_i^+ \cap A_d} \frac{1}{c_{ag}}} + \sum_{a \in A_i^+} p_a^{dg} \cdot w_{N_a^+}^{dg}, \tag{2.5}$$

$$p_a^{dg} = B\left(a \in A_i^+ \cap A_d\right) \cdot \frac{\dfrac{1}{c_{ag}}}{\displaystyle\sum_{b \in A_i^+ \cap A_d} \frac{1}{c_{bg}}}, \quad \forall a \in A_i^+. \tag{2.6}$$

This formulation results from the classical optimal strategy model (Spiess and Florian, 1989) under the hypothesis that the cost of each waiting arc exiting the stop is proportional to the headway of the corresponding line.

Finally, for a placing node $i \in P$ it is:

$$w_i^{dg} = 0 + \sum_{a \in A_i^+} p_a^{dg} \cdot w_{N_a^+}^{dg} \, , \qquad (2.7)$$

$$p_a^{dg} = \pi_{ag}, \quad \forall a \in A_i^+ \, . \qquad (2.8)$$

Both the placing probabilities and the waiting arc costs depend on passengers' flows; the form of these relations will be presented in the congestion model.

2.3. Flow propagation model

The classical flow propagation model consists in loading the demand flows towards each destination, according to the arc conditional probabilities resulting from the route choice model as follows:

$$q_i^{dg} = d_{idg} + \sum_{b \in A_i^-} y_b^{dg} \, , \qquad (2.9)$$

$$y_a^{dg} = q_i^{dg} \cdot p_a^{dg}, \quad \forall a \in A_i^+ \, . \qquad (2.10)$$

The first equation yields the flow exiting from each node as the sum of the travel demand and the entering arc flows. The second one propagates the node flow on the arcs of its forward star.

In the following, to improve convergence to equilibrium we adopt gradient projection when finding the search direction y_a^{dg} in the propagation model of used standard nodes, as follows:

$$p_a^{dg} = Max\left(0, \frac{q_a^{dg}}{\sum\limits_{b \in A_i^+} q_b^{dg}} + \frac{1}{2} \cdot \frac{w_i^{dg}}{c_{ag} + w_{N_a^+}^{dg}} - \frac{1}{2} \right), \quad \forall a \in A_i^+ - s_i^{dg} \, , \qquad (2.11)$$

$$p_a^{dg} = 1 - \sum_{b \in A_i^+} p_b^{dg}, \quad a = s_i^{dg} \, . \qquad (2.12)$$

In the proposed implementation of the gradient projection schema, the probability shift of each non-minimal local alternative is negative and equals to half of the complement to one of the ratio between the minimum cost from node i to destination d and the cost to reach d using arc a. When these two costs coincide, the shift is null as the local flows satisfy the conditions for deterministic equilibrium. The resulting probabilities must anyhow be non-negative. Finally, the probability of the node successive arc acquires all the shifts subtracted to the non-minimal alternatives, but not necessarily all their probabilities, as it would happen instead in the classical All-Or-Nothing assignment.

For stops, placing and unused nodes we stick to the probabilities of the route choice model.

Note that gradient projection requires the previous values of the conditional probabilities that are obtained from the destination specific flows q_a^{dg}.

To measure convergence, we introduce the following version of the relative gap:

$$\rho = \frac{\sum\limits_{i \in N-(S+P)} \sum\limits_{a \in A_i^+} q_a^{dg} \cdot \left(c_{ag} + w_{N_a^+}^{dg} - w_i^{dg} \right) + \sum\limits_{i \in S+P} w_i^{dg} \cdot \sum\limits_{a \in A_i^+} \left\| q_a^{dg} - p_a^{dg} \cdot \sum\limits_{b \in A_i^+} q_b^{dg} \right\|}{\sum\limits_{i \in N-(S+P)} \sum\limits_{a \in A_i^+} q_a^{dg} \cdot \left(c_{ag} + w_{N_a^+}^{dg} + w_i^{dg} \right) + \sum\limits_{i \in S+P} w_i^{dg} \cdot \sum\limits_{a \in A_i^+} \left(q_a^{dg} + p_a^{dg} \cdot \sum\limits_{b \in A_i^+} q_b^{dg} \right)} \, ; \qquad (2.13)$$

because at stops and placing nodes no decision is actually taken, a flow difference is considered instead of the gap with respect to the minimum cost.

The averaging model which is required to guarantee the convergence of the gradient projection method is like the Method of Successive Averages, where k^{bad} is the number of previous iterations that did not provide a sufficient improvement of the relative gap:

$$\alpha = \frac{1}{1 + k^{bad}} \ , \tag{2.14}$$

$$q_a^{dg} = y_a^{dg} \cdot \alpha + q_a^{dg}\left(1 - \alpha\right). \tag{2.15}$$

The aggregation model consists in a simple sum over all destinations:

$$q_{ag} = \sum_{d \in Z} q_a^{dg} \ . \tag{2.16}$$

3. Supply model

3.1. Transit network

The proposed supply model simulates the line services and the pedestrian network that connects the stops for access, egress and transfer, while the extension to multimodal transport and shared mobility (like park & ride, car sharing and ride pooling) is out of the scope of this paper.

Hence, the focus is on public transport, which is perceived from passengers as one single mode composed of a variety of physical systems (such as metro, train, bus, and sidewalks).

The directed graph *(N, A)* adopted to model the transit network is articulated in sub-networks, consisting of separate layers of arcs and nodes that represent specific transport systems, where each arc represents a specific trip phases. Figure 3.1 depicts the graph elements of the generic stop $s \in S_\ell$ of line $\ell \in L$.
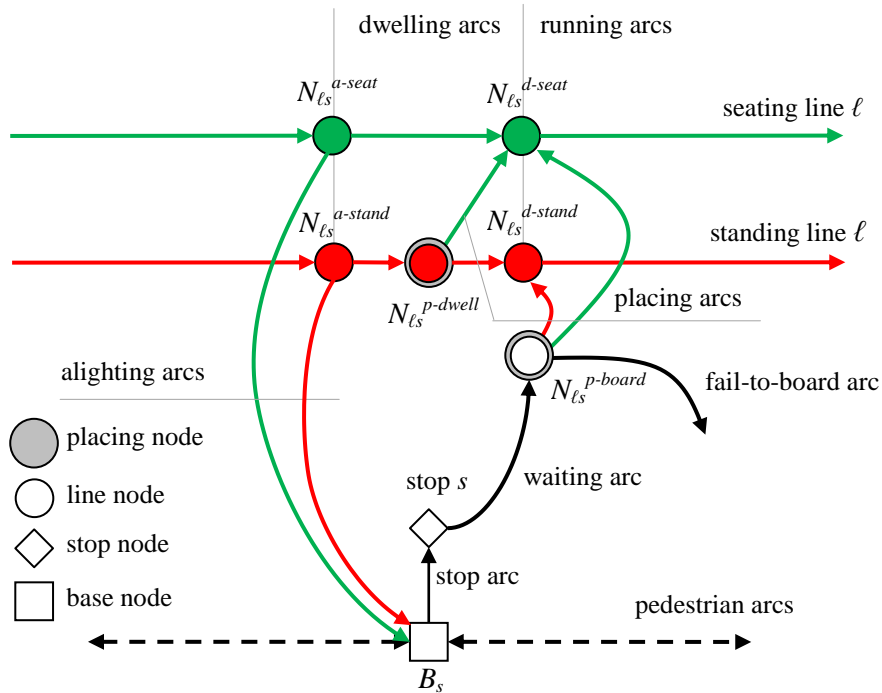

Figure 3.1. Topology of the transit network for a generic line $\ell \in L$ and its generic stop $s \in S_\ell$

More specifically, the transit network is composed of:
- the pedestrian network, which is the portion of the base (infrastructural) network where passengers can walk, including the connectors to the zone centroids;
- the network of each line $\ell \in L$, which represents riding, dwelling, boarding and alighting vehicles, and it is duplicated into seating and standing to better reproduce these two different comfort conditions;
- stop arcs that connect the pedestrian network to the line networks.

Thus, the graph is composed of the following types of nodes and links:
- base nodes N_{base}, including the zone centroids Z and the base stops B_s,
- stop nodes S,
- line nodes N_ℓ, including the placing nodes P,
- running arcs A_{run},

- base arcs A_{base},
- stop arcs A_{stop},
- waiting arcs A_{wait},
- placing arcs A_{place},
- alighting arcs A_{alight},
- dwelling arcs A_{dwell}.

Passengers are subjected to unpredictable events at particular nodes. More precisely, at stops passengers wait for service and are subject to random vehicle arrivals; at placing nodes, mingling passengers try to board the vehicle and seat but are subject to a failing probability, which depends on the ratio between available capacity and request flow.

A key innovation of this model, with respect to the transit assignment models available in the most used software packages for transport network simulation, is that different congestion phenomena can be simulated through the introduction of non-separable arc cost functions, which will be presented in the following.

3.2. Arc costs

The *generalized cost* c_{ag} of arc $a \in A$ for user of class $g \in G$ is the sum of a *non-temporal cost* χ_{ag}, which is not affected by congestion, and a temporal cost, given by the *value of time* γ_{ag} multiplied by the *travel time* t_a :

$$c_{ag} = \chi_{ag} + \gamma_{ag} \cdot t_a . \tag{3.1}$$

The non-temporal cost is in general a weighted sum of *arc attributes* n_a^k (transit fares, link length, stop and on-board facilities, etc.) multiplied by *attribute coefficients* γ_g^k that express the route choice preferences of the class:

$$\chi_{ag} = \sum_k n_a^k \cdot \gamma_g^k . \tag{3.2}$$

A random utility model for route choice should be calibrated to set the values of these coefficients.

The specific value of time of each arc can be computed by multiplying the *base value of time* γ_g^{vot} of the class for different *discomfort coefficients* λ_{ag} which change according to the arc type and flows:

$$\gamma_{ag} = \gamma_g^{vot} \cdot \lambda_{ag} . \tag{3.3}$$

Null cost is assumed for stop and placing arcs.

3.3. Pedestrian arcs

Pedestrian arcs represent passengers walking on sidewalks.

The travel time of the generic base arc $a \in A_{base}$ is usually given by the link length divided by the walking speed. When pedestrian congestion is an issue a BPR function like the following can be adopted:

$$t_a = \frac{l_a}{v^{walk}} \cdot \left(1 + \alpha^{walk} \cdot \left(\frac{q_a}{k^{walk} \cdot w_a} \right)^{\beta^{walk}} \right), \qquad a \in A_{base}, \tag{3.4}$$

where:
- l_a is the link length,
- v^{walk} is the walking speed,
- k^{walk} is the capacity per unit width of sidewalks,
- w_a is the width of the sidewalk (if null, time is assumed as infinite,
- α^{walk} and β^{walk} are the BPR parameters.

The non-temporal cost can be obtained by considering among the relevant arc attributes the number of shops and other points of interest, the slope, the presence of public lights and other safety and security features. However, in most transit applications pedestrian flows are not a critical outcome, thus

the above qualitative characteristics are often disregarded despite their relevance (i,e, the non-temporal cost is simply assumed equal to the travel time).

Finally, a constant discomfort coefficient λ_g^{walk} is considered.

3.4. Running arcs

The generic running arc $a \in A_{run}$ represents passengers sitting or standing on-board of a vehicle of line $\ell \in L$ travelling from one stop $s \in S_\ell$ to the successive stop $s+1$.

For what concerns travel time, in case a timetable exists for line ℓ then it is computed as the average, over all runs $r \in R_\ell$, of the difference between the head stop arrival time and the tail stop departure time:

$$t_a = \frac{\sum_{r \in R_\ell} \tau_{r\,s+1\ell} - \vartheta_{rs\ell}}{|R_\ell|}, \qquad a = \begin{cases} \left(N_{\ell s}^{d-seat}, N_{\ell s+1}^{a-seat} \right) \\ \left(N_{\ell s}^{d-stand}, N_{\ell s+1}^{a-stand} \right) \end{cases} ., \tag{3.5}$$

where:

- $\tau_{rs\ell}$ is the arrival time of line ℓ at stop s,
- $\theta_{rs\ell}$ is the arrival time of line ℓ at stop s.

The non-temporal cost is obtained multiplying the length of the link l_a by the kilometric fee c_ℓ^{kfee} for the corresponding line and by a possible fee multiplier γ_g^{mfee} for the class. It may further take into account a large set of attributes that describe the quality of service onboard the vehicle, e.g. air conditioning, seat ergonometry, passenger entertainment and information, etc.

Otherwise, the commercial speed v_b^{line} of the support edges $b \in A_{\ell s}$ of the line segment is considered:

$$t_a = \sum_{b \in A_{\ell s}} \frac{l_b}{v_b^{line}} \quad , \qquad a = \begin{cases} \left(N_{\ell s}^{d-seat}, N_{\ell s+1}^{a-seat} \right) \\ \left(N_{\ell s}^{d-stand}, N_{\ell s+1}^{a-stand} \right) \end{cases} .. \tag{3.6}$$

The non-temporal cost or the running arc may consider a large set of attributes that describe the quality of service onboard the vehicle, e.g. air conditioning, seat ergonometry, passenger entertainment and information, etc.

Finally, the discomfort coefficient assumes two different values λ_g^{stand} and λ_g^{seat} depending on the arc type under consideration.

3.5. Dwelling arc

The generic dwelling arc $a \in A_{dwell}$ represents a vehicle of line $\ell \in L$ still at stop $s \in S_\ell$ to allow passengers alighting and boarding. Dwelling is a sequence of operations: 1) doors opening after the vehicle is safely positioned at the stop; 2) passenger alighting and boarding; 3) doors remaining open without passenger flow (if necessary); 4) doors closing and safety control before vehicle departure.

In case a timetable exists for line ℓ, then the travel time is computed as the average, over all runs $r \in R_\ell$, of the difference between the departure time and the arrival time at that stop:

$$t_a = \frac{\sum_{r \in R_\ell} \vartheta_{rs\ell} - \tau_{rs\ell}}{|R_\ell|}, \qquad a = \begin{cases} \left(N_{\ell s}^{a-seat}, N_{\ell s}^{d-seat} \right) \\ \left(N_{\ell s}^{a-stand}, N_{\ell s}^{p-dwell} \right) \end{cases} . \tag{3.7}$$

Otherwise, we assume the following. The first and last operations are independent of passenger flows and require a fixed door manoeuvre time t_ℓ^{doors}. The second operation is associated to the passenger alighting and boarding time t_a^{ab}. The third operation is done only if the minimum dwelling time t_ℓ^{dwell} is not yet passed. Thus, it is:

$$t_a = Max\left(t_\ell^{dwell}, t_\ell^{doors} + t_a^{ab} \right), \qquad a = \begin{cases} \left(N_{\ell s}^{a-seat}, N_{\ell s}^{d-seat} \right) \\ \left(N_{\ell s}^{a-stand}, N_{\ell s}^{p-dwell} \right) \end{cases} . \tag{3.8}$$

The time required for passenger alighting and boarding is computed differently according to whether doors are dedicated or not:

$$t_a^{ab} = h_{\ell s} \cdot \begin{cases} \dfrac{q_d + q_e}{k_\ell^{alight}} + \dfrac{q_b}{k_\ell^{board}}, & \text{not dedicated doors} \\[3mm] Max\left(\dfrac{q_d + q_e}{k_\ell^{alight}}, \dfrac{q_b}{k_\ell^{board}} \right), & \text{dedicated doors} \end{cases}, \qquad \begin{aligned} a &= \begin{cases} \left(N_{\ell s}^{a-seat}, N_{\ell s}^{d-seat} \right) \\[2mm] \left(N_{\ell s}^{a-stand}, N_{\ell s}^{p-dwell} \right) \end{cases} \\[2mm] b &= \left(s, N_{\ell s}^{p-board} \right) \\[2mm] d &= \left(N_{\ell s}^{a-seat}, B_s \right) \\[2mm] e &= \left(N_{\ell s}^{a-stand}, B_s \right) \end{aligned}, \qquad (3.9)$$

where:
- $h_{\ell s}$ is the expected headway of line $\ell \in L$ at stop $s \in S_\ell$;
- k_ℓ^{alight} and k_ℓ^{board} are, respectively, the alighting and boarding capacity of vehicles serving line ℓ.

Moreover, this time may be multiplied by a BPR factor to represent overcrowding congestion as explained later.

For simplicity, the non-temporal cost of dwelling arcs is null, while the discomfort coefficient is just the base value of time.

3.6. Alighting arcs

The alighting arc $a \in A_{alight}$ represents passengers getting off a vehicle of line $\ell \in L$ at stop $s \in S_\ell$.

The travel time of alighting arcs is assumed null.

The non-temporal cost is given by the transfer cost c_g^{tran} for the user class $g \in G$ and represents different aspects of transfer disutility, not necessarily connected with a measurable delay (e.g. psychological stress of changing line, additional travel time variance induced by random vehicle departures, etc.).

A specific boarding fee c_ℓ^{bfee} for line ℓ multiplied by the fee multiplier γ_g^{mfee} for the class may be added to the non-temporal cost.

3.7. Waiting arcs

The waiting arc $a \in A_{wait}$ represents passengers waiting to board a vehicle of line $\ell \in L$ at stop $s \in S_\ell$.

In case of frequency-based services, the free flow time is the expected waiting time:

$$t_a = \frac{h_{\ell s}}{2} \cdot \left(1 + \sigma_{\ell s}^2 \right), \qquad a = \left(s, N_{\ell s}^{p-board} \right), \qquad (3.10)$$

where:
- $\sigma_{\ell s}$ is the variation coefficient of the headway distribution, that represents the service irregularity of line $\ell \in L$ at stop $s \in S_\ell$ and is given by the ratio between the headway standard deviation and its expected value $h_{\ell s}$.

If actual realizations of departure times are available (e.g. from AVM), it is possible to compute the above as follows:

$$h_{\ell s} = \frac{\sum\limits_{r>1 \in R_\ell} \vartheta_{r+1 s \ell} - \vartheta_{r s \ell}}{|R_\ell| - 1}, \qquad (3.11)$$

$$\sigma_{\ell s}^2 = \frac{1}{h_{\ell s}^2} \cdot \frac{\sum\limits_{r>1 \in R_\ell} \left(\vartheta_{r+1 s \ell} - \vartheta_{r s \ell} - h_{\ell s} \right)^2}{|R_\ell| - 1}. \qquad (3.12)$$

In case of schedule-based services, the free flow time is instead given by the most convenient option between waiting at the stop for half of the expected headway on average, or only for a boarding time t_ℓ^{board} of line $\ell \in L$ associated to a safety margin:

$$t_a = Min\left(\frac{h_{\ell s}}{2}, t_\ell^{board}\right), \qquad a = \left(s, N_{\ell s}^{p-board}\right). \tag{3.13}$$

To represent queuing congestion, the free flow waiting time may be multiplied by a BPR factor as it will be explained in section 3.11.

The non-temporal cost of waiting arcs may take into account a large set of attributes that describe the quality of service at the stop, e.g. shelter, seats, passenger information, shops, etc.

A discomfort coefficient λ_g^{wait} is also considered, which may be multiplied by a BPR factor in case of overcrowding congestion as it will be explained in section 3.9.

3.8. Placing arcs

Passengers travelling on transit services experience more discomfort by standing versus sitting. Passengers are subjected to the random process of finding a seat on-board of line $\ell \in L$ at stop $s \in S_\ell$ at two different placing nodes of the graph: $N_{\ell s}^{p-dwell}$ and $N_{\ell s}^{p-board}$.

Passengers arriving at a stop with a seat have guaranteed a seat for the next line segment, so that they either alight or remain seated. Passengers arriving at a stop standing who do not alight have priority over the passengers newly boarding, in the sense that they have an earlier chance to occupy any seat that might become vacant thanks to alighting passengers.

The probability of seating depends on the ratio between the remaining seats and the passengers attempting to sit. The remaining seats are in turn obtained by subtracting to the nominal vehicle capacity of seats k_ℓ^{seat} the number of passengers already seated. Hence, the placing probability is anyhow bounded between 0 and 1.

The probability of standing can be computed as the complement to one of the other placing probabilities. At node $N_{\ell s}^{p-board}$ these may include the fail-to board probability. However, for what concerns the expected cost of reaching the destination, the fail-to-board arc is not considered in the weighted average, while the cost of failing is directly added to the waiting arc (see section 3.11).

At node $N_{\ell s}^{p-dwell}$ the sitting probability is:

$$\pi_a^{seat} = Mid\left(0, \frac{\dfrac{k_\ell^{seat}}{h_{\ell s}} - q_d}{q_b}, 1\right), \qquad \begin{aligned} a &= \left(N_{\ell s}^{p-dwell}, N_{\ell s}^{d-seat}\right) \\ b &= \left(N_{\ell s}^{a-stand}, N_{\ell s}^{p-dwell}\right). \\ d &= \left(N_{\ell s}^{a-seat}, N_{\ell s}^{d-seat}\right) \end{aligned} \tag{3.14}$$

At node $N_{\ell s}^{p-board}$ the sitting probability is:

$$\pi_a^{seat} = Mid\left(0, \frac{\dfrac{k_\ell^{seat}}{h_{\ell s}} - q_d - q_e}{q_b}, 1\right), \qquad \begin{aligned} a &= \left(N_{\ell s}^{p-board}, N_{\ell s}^{d-seat}\right) \\ b &= \left(s, N_{\ell s}^{p-board}\right) \\ d &= \left(N_{\ell s}^{a-seat}, N_{\ell s}^{d-seat}\right) \\ e &= \left(N_{\ell s}^{p-dwell}, N_{\ell s}^{d-seat}\right) \end{aligned} \tag{3.15}$$

The node expected cost is the average cost of sitting and standing, weighted by the sitting and fail-to-sit probability. In turn, the cost of standing includes the possibility of sitting at next stops.

The alighting decision is not predetermined anymore. Indeed, passengers who have obtained a seat might prefer to transfer later, whereas standing passengers are more likely to transfer earlier.

3.9. Overcrowding congestion

The more passengers standing on crowded vehicles and platforms are physically packed, the more they perceive travelling as uncomfortable and stressful. For medium densities, crowding discomfort is due to the close physical distance with other passengers; for higher densities, it is due to contact and pressure of other passengers. The discomfort is then assumed to increase exponentially with the saturation rate, i.e. the number of passengers divided by the capacity.

Only for passengers standing on-board a vehicle (seated passengers are not affected), the discomfort coefficient of running arc $a \in A_{run}$ of line $\ell \in L$ at stop $s \in S_\ell$ is multiplied by the following BPR-type factor:

$$1 + \alpha_g^{crowd} \cdot \left(\frac{q_a \cdot h_{\ell s}}{k_\ell^{stand}} \right)^{\beta^{crowd}} , a = \left(N_{\ell s}^{d-stand}, N_{\ell s+1}^{a-stand} \right), \tag{3.16}$$

where:
- k_ℓ^{stand} is the standing capacity of vehicles serving line ℓ;
- α^{crowd} and β^{crowd} are the BPR parameters (typical values are respectively 1 and 2).

Similarly, the discomfort coefficient of waiting arc $a \in A_{wait}$ for all passengers standing at stop $s \in S_\ell$ is multiplied by the following BPR-type factor:

$$1 + \alpha^{crowd} \cdot \left(\frac{\sum_{b \in A_s^+} q_b \cdot t_b}{k_s} \right)^{\beta^{crowd}} , \quad a = \left(s, N_{\ell s}^{p-board} \right), \tag{3.17}$$

where:
- k_s is the platform capacity of stop s,
- the sum of products between the passenger flow of each waiting arc exiting from the stop q_b and its time t_b yields the average number of passengers waiting at the stop s.

It can be observed that the crowding discomfort for onboard passengers depends only on the flow of standing passengers (separable congestion), while the crowding discomfort for passengers at stop depends on several arc flows (non-separable).

3.10. Dwelling delay

The nominal capacity of doors, introduced in section 3.5, can be reduced by on-board overcrowding due to the difficulty of moving inside the carrier, so that boarding/alighting passengers take more time to access/egress the vehicle.

To simulate this effect, the time required for passenger alighting and boarding t_a^{ab} of the generic dwelling arc $a \in A_{dwell}$ is multiplied by the following BPR factor causing an additional dwelling delay:

$$1 + \alpha^{dwell} \cdot \left(\frac{q_d \cdot h_{\ell s}}{k_\ell^{stand}} \right)^{\beta^{dwell}} , \quad \begin{aligned} a &= \begin{cases} \left(N_{\ell s}^{a-seat}, N_{\ell s}^{d-seat} \right) \\ \left(N_{\ell s}^{a-stand}, N_{\ell s}^{p-dwell} \right), \end{cases} \\ d &= \left(N_{\ell s}^{a-stand}, N_{\ell s}^{p-dwell} \right) \end{aligned} \tag{3.18}$$

where α^{dwell} and β^{dwell} are the BPR parameters (typical values are respectively 1 and 2).

3.11. Queuing congestion

When overcrowding is severe, and the crush capacity is reached on-board, no further passenger can get on the vehicle and stand. Then, an over-saturation queue of passengers waiting at the stop is formed, increasing the expected waiting time. Clearly, this phenomenon does not affect the passengers that are already on-board, but only those willing to board.

The queuing process may follow different regimes depending on the stop layout. In stations with large platforms passengers do mingling, while at urban bus stops they tend to respect the boarding priority of anybody who has arrived before them. In any case, the additional wait time due to the lack of space on-board increases not only with the number of passengers wishing to board but also with the number of dwelling passengers that are already on-board.

Two main modelling approaches may be adopted to represent queuing congestion:
- soft capacity constraints, through effective frequencies;
- strict capacity constraints, through fail-to-board probabilities.

In the first case, the vehicle capacity can be exceeded by the number of on-board passengers. Congestion affects the cost pattern through the so-called *effective frequency* (De Cea and Fernandez, 1993), inducing additional impedance on waiting arcs by decreasing the nominal frequency of the line.

Despite the route choice model will indirectly tend to lower the on-board flow exceeding the line capacity, relevant capacity violations can result at the equilibrium when no alternative route is available. In our model, capacity violations are allowed only for standing passengers, as it is assumed that they can squeeze, contrarily to sitting passengers. The method may result in travel times that are unrealistically high as static assignment models are not able to reproduce the capacity accumulation of the network.

The fundamental idea behind this approach is that, when passengers mingle at stops, the probability to succeed in boarding the approaching vehicle is related to the standing saturation rate of the next line segment (running arc d), where the waiting flow and the dwelling flow merge. Therefore, the effective frequency is here implemented by multiplying the free flow time of the generic waiting arc $a \in A_{wait}$ by the following BPR factor:

$$1 + \alpha^{queue} \cdot \left(\frac{q_d \cdot h_{\ell s}}{k_\ell^{stand}} \right)^{\beta^{queue}} \quad , \quad \begin{aligned} a &= \left(s, N_{\ell s}^{p-board} \right) \\ d &= \left(N_{\ell s}^{a-stand}, N_{\ell s}^{p-dwell} \right) \end{aligned} \quad , \tag{3.19}$$

where α^{queue} and β^{queue} are the BPR parameters (typical values are respectively 1 and 4).

In the second case (i.e. strict capacity constraints), the vehicle capacity will never be exceeded by the number of on-board passengers. The fail-to-board probability method (Kurauchi *et al.*, 2003) removes the flow in excess from the waiting arc and ideally injects it in the following temporal layer of a quasi-dynamic assignment (e.g. next hour) through the introduction of specific dummy arcs.

The fail-to-board probability approach is based on the idea that, when mingling queues of passengers occur at the stop, the probability to succeed in boarding the next approaching vehicle depends on the ratio between the remaining vehicle capacity and the passengers wishing to board (waiting arc b). The remaining capacity is in turn obtained by subtracting to the nominal capacities for seating and standing the number of dwelling passengers seating and standing (dwelling arcs e and d, respectively).

The placing probability of the generic failing arc $a \in A_{fail}$ that represents passengers failing to board a vehicle of line $\ell \in L$ at stop $s \in S_\ell$ is then given by:

$$\pi_a^{fail} = 1 - Mid \left(0, \frac{\dfrac{k_\ell^{stand} + k_\ell^{seat}}{h_{\ell s}} - q_d - q_e}{q_b}, 1 \right) \quad , \quad \begin{aligned} a &= \left(N_{\ell s}^{p-board}, N_{\ell s}^{fail} \right) \\ b &= \left(s, N_{\ell s}^{p-board} \right) \\ d &= \left(N_{\ell s}^{a-seat}, N_{\ell s}^{d-seat} \right) \\ e &= \left(N_{\ell s}^{a-stand}, N_{\ell s}^{p-dwell} \right) \end{aligned} \quad . \tag{3.20}$$

Hence, when there is not enough remaining capacity on board, some travellers will fail to board. In practice, as a static assignment is performed, the passengers who failed to board are eliminated from the model. This is a relevant drawback for planning purposes; therefore, the model has mainly an operational validity.

Furthermore, all waiting passengers suffer from a supplementary time c_{bg}^{fail} due to the risk of failing-to-board, which is additional to the temporal cost of the waiting arc $b \in A_{wait}$ and is proportional to:

- the value of waiting time
- the risk-averseness coefficient λ_g^{risk} of the class towards abnormal delays
- the probability of failing π_a^{fail}
- the additional waiting time in case of failing, which is given by the headway multiplied by the expected number of vehicle arrivals before successful boarding

$$c_{bg}^{fail} = \gamma_g^{vot} \cdot \lambda_g^{wait} \cdot \lambda_g^{risk} \cdot \pi_a^{fail} \cdot \frac{h_{\ell s}}{1 - \pi_a^{fail}} \quad , \quad \begin{aligned} a &= \left(N_{\ell s}^{p-board}, N_{\ell s}^{fail} \right) \\ b &= \left(s, N_{\ell s}^{p-board} \right) \end{aligned} \quad . \tag{3.21}$$

Thus, the failing cost tends to infinity as the fail to board probability goes to one. The number of passengers who will accept the risk of failing is a result of the equilibrium mechanism.

4. Strategies and hyperpaths

4.1. Route alternatives on transit networks

In transport modelling, a strategy can be defined as the plan users adopt to reach their destination.

In presence of unpredictable events, such as random departures of vehicles from stops and failing probabilities at placing nodes, it may be not convenient or simply impossible to choose and follow a specific path on the graph to accomplish a given trip. For example, it could be better to board a slower line that is arriving earlier than to wait for a faster line that will arrive later, where slow and fast is intended in term of expected travel time to reach the destination once boarded the line. On the other hand, an unlucky passenger who was not able to sit, shall necessarily ride the next line segment standing onboard the vehicle.

Thus, passengers must cope with a more complex definition of route to travel on transit networks, as they have to identify how to follow their trip toward the destination for each local alternative of random diversions they encounter en-trip. Placing and stops are nodes where passengers acquire information about variables that are preventively seen as random unknowns, and on this base can make en-route decisions on how to proceed toward the destination.

Among all the available strategies, it is assumed that passengers choose one with minimum expected cost before starting the trip. For example, passengers choose at home the set of attractive lines they can board at a given stop (this is a strategic choice). Once at the stop, they will discover which line to take as the outcome of an event, i.e. which is the first vehicle of an attractive line approaching the stop. Clearly, the journey will follow a different path depending on which event has happened. The case of placing nodes with sitting probabilities is slightly different, because the passenger cannot make any strategic choice there, and must simply accept the outcome of the event hoping to be lucky enough to find a seat; in this case the route choice is made at earlier nodes.

4.2. Hyperpath topology

From a topological point of view, strategies on transit networks are formalized with hyperarcs and hyperpaths.

By definition, a hyperarc $\breve{a}$ is a non-empty set of arcs exiting from a diversion node $i \in D \subseteq N$, i.e. a subset of its forward star A_i^+. Note that not all combinations of such exiting arcs make a valid hyperarc, i.e. the set of hyperarcs H must be specified. Each branch $a \in \breve{a}$ of the hyperarc is characterized by the diversion probability $p_{a|\breve{a}}$ and by a conditional cost $c_{a|\breve{a}}{}^g$.

The generic hyperpath k is a bush of arcs, i.e. an acyclic sub graph that connects an origin to a destination with one successor arc for each one of its nodes and one hyperarc for each one of its diversions, except for the destination which has none. Its cost is defined as the sum of the arc costs and of the hyperarc branch costs, multiplied by the probability of using them.

4.3. Hyperarcs at stops and placing nodes

In the proposed model two types of diversion nodes have been introduced: stop nodes and placing nodes.

The hyperarcs exiting from a stop node are all possible sets of waiting arcs, i.e. all the combinations of available lines. Among such waiting hyperarcs, the passenger directed towards a given destination can choose the so called attractive set of lines. For a given waiting hyperarc, the diversion probability of each branch is given by the frequency divided by the combined frequency, while the conditional cost is equal to the inverse of the combined frequency. Because the line frequency is the inverse of the headway and the latter is proportional to the waiting arc cost presented in section 3.7, the following is obtained:

$$p_{a|\breve{a}} = \frac{\dfrac{1}{c_{ag}}}{\displaystyle\sum_{b \in \breve{a}} \dfrac{1}{c_{bg}}}, \tag{4.1}$$

$$c_{a|\breve{a}}^{\,g} = \frac{1}{\displaystyle\sum_{b \in \breve{a}} \dfrac{1}{c_{bg}}}. \tag{4.2}$$

The above model is valid only if the headways are distributed as independent exponentials.

There is only one hyperarc exiting from each placing node that is given by the set of all placing and failing arcs. Noteworthy, for seating hyperarcs there is no choice to be made. The diversion probabilities are given by the placing probabilities while the conditional cost is null.

4.4. Shortest hyperpaths and absorbing cycles

The route choice model introduced in section 2.2 is a sequential one based on implicit enumeration of routes. Users reach their destination through a sequence of local choices at nodes and probabilistic events at diversions, where the local alternatives are the arcs of the forward star. The conditional probabilities depend on the arc costs and on the placing probabilities provided by the congestion model. There was no need of explicitly introducing hyperarcs and hyperpaths, but this simplification came together with two relevant assumption: only efficient arcs are used and the attractive set at stops is implicitly predetermined by the given topological order.

In this section we extend the model to remove these two limiting assumptions.

The minimum strategy cost from each node $i \in N$ to destination $d \in Z$ can be obtained through an extension of Dijkstra algorithm (Dijkstra, 1959) that computes a shortest hypertree which repeatedly applies the following Bellman relation until no improvement is possible:

$$w_i^{dg} \leftarrow Min \begin{cases} \sum_{a \in \breve{a}} p_{a|\breve{a}} \cdot \left(c_{a|\breve{a}}^g + w_{N_a^+}^{dg} \right), \forall \breve{a} \subseteq A_i^+ \cap H \quad i \in D \\ c_{ag} + w_{N_a^+}^{dg}, \forall a \in A_i^+ \quad i \in N - D \end{cases}. \tag{4.3}$$

The method starts from infinite node costs except for the destination which has null cost.

When a topological order is given and only efficient arcs are used, the application of the above schema is trivial. Otherwise, head nodes can be extracted from a list in order of cost and their backward star is visited; in case of improvement, the updated tail node is inserted in the list. This Dijkstra schema holds if we apply the first of (4.3) only to hyperarcs whose other heads have been already visited. The resulting algorithm is label setting (nodes are visited only once) if the updated tail node has a higher cost than the extracted head node.

In the case of stop nodes, this is the case. The attractive set is computed with a Greedy algorithm, which exploits the order of lines in terms of remaining costs to reach the destination once boarded. Lines are added to the attractive set (initially empty) in increasing order of remaining cost. The algorithm stops when the remaining cost of the next line is higher than the current value of the expected cost.

As congestion increases, more (and hence slower) lines are included in the attractive set. Moreover, if all lines are congested, some passengers would rather walk than continue to wait.

Casting the model 4.1, 4.2 into equation 4.4 we reobtain equation 2.5.

Unfortunately, in the case of placing nodes this is not the case, because the cost of the updated placing node is a weighted average of the extracted standing node which has a higher cost than the other seating node. This compromises the label setting approach of the Dijkstra algorithm. Indeed, nodes with lower cost could be extracted after nodes with a higher cost, so that a node already extracted can be further optimized (and can be visited more than once).

As a consequence of the presence of head costs higher than tail ones, the optimal strategy can involve so-called absorbing cycles. For example, an unlucky boarding passenger unable to seat, who then alights at next stop may walk back to wait again for the line at the previous stop, thus gaining another chance of sitting on-board). This is not only inconsistent with the definition of hyperpath as a bush (which is by definition acyclic), but also presumes a non-realistic behaviour.

To avoid this kind of paradoxes, we accept the label correcting except for alighting arcs, where a label setting approach is imposed. This ensures the acyclicity of the solution, although it may lead to non-optimal hypertrees. In practice we did not notice difficulties in convergence to equilibrium.

4.5. Gradient projection and implicit enumeration of hyperarc

The drawback of fixed-point problems with respect to more classical optimization models (e.g. where the objective function is the sum of cost integrals) is the lack of rapidly convergent algorithms, which prevents precise calculations of the equilibrium.

The Method of Successive Averages (MSA) is one of the most used approaches to solve fixed-point problems and is shown to converge to equilibrium under Blum theorem conditions (see Cascetta 2009, Appendix A). MSA is a very simple method: the new iterate is the average of all network loading results of previous iterations (usually, an All-Or-Nothing assignment to shortest routes implements each network loading). But MSA does not give high convergence precision. It usually tails badly after few iterations because the step size becomes smaller at each iteration.

To overcome MSA drawbacks, the solution algorithm implements a reduced gradient projection (RGP) method over implicit hyperarcs. RGP is a variant of Gradient Projection (GP) methods, which consist in iterative optimization algorithms for constrained functions that project the search direction on the active constraints (Rosen, 1960).

The implemented RGP is an innovative one, as it performs convergence search over arcs, avoiding the enumeration of hyperarcs (RGP on implicit hyperarcs). The implicit definition of hyperarcs is an important goal achieved, as enumerating them would be costly in terms of computational time. Indeed, transit networks usually consists in large-scale instances and the total number of possible hyperarcs is usually huge (e.g. considering a small network composed of three stops and three lines serving those stops, there are already 16 hyperarcs: 7 waiting hyperarcs and 9 seating hyperarcs).

In general, GP algorithms for UE computation find the optimal set of alternatives and flows, solving the convex cost optimization problem (Gentile, 2016), where alternatives can be either hyperpaths connecting the OD pairs or arcs exiting a node towards the destinations.

5. Conclusions

5.1. Main achievements

Governments continue to invest significantly on public transport. hence anticipating its development and performances will continue to play a central role in the future in urban planning.

The ability of public transport systems to attract demand from private transport represents a key success factor. But congestion and discomfort may negatively affect user experience, and, in turn, passenger choices are often influenced by their severity.

Careful strategic and tactical planning decisions supported by transport modelling solutions can help increasing the level of service, hence public transport ridership. It is of paramount importance that models comprehensively represent all the relevant phenomena. Different types of models for public transport have been proposed and implemented in the past 50 years. However, examples of practical implementations that include passenger congestion phenomena, despite their importance, are still very limited. Moreover, the existing simulation models require high computational burden.

Through a practical implementation of the equilibrium model and a detailed discussion of its supply side, this paper presents a deterministic static assignment method using a frequency-based approach for the simulation of transit networks, which is capable of representing several congestion phenomena such as overcrowding congestion, queuing congestion, and dwelling delay. Moreover, with an efficient gradient projection algorithm we can obtain User Equilibrium for large-scale networks with high computation performance. In addition, the model represents other phenomena, such as the availability of seats (both for boarding passengers and dwelling passengers) and the waiting process at stops.

Further, the adopted assignment algorithm computes a strategy-based user equilibrium with implicit hyperpath enumeration by solving a fixed-point formulation, through a reduced gradient projection method, which highly improves convergence performance. The implicit definition of hyperarcs is another important goal achieved, as this would add significant computational time.

Finally, the model has been implemented in .net which interfaces with PTV Visum as I/O source. Simulations and validation are focused on the analysis of the various transit phenomena which are tested singularly. The numerical results will be presented in a following paper.

5.2. Further developments

Congestion and overcrowding are not only determined by wrong planning decisions but also by the lack of operational and real-time planning. And yet, a system which can predict the network response to incidents and propose a solution to keep congestion levels acceptable does not yet exist.

Indeed, to mitigate congestion events and keep the level of service high, public transport operators should employ a system able to predict congestion development and guide users' choices. However, a software fast enough to forecast real-time passengers' congestion in public transport networks while considering a wide range of congestion phenomena, does not exist yet.

The model presented in this paper is not suitable for real-time simulation and incident management, as it performs a static assignment which implies steady state setting (constant flows and performances during the assignment period), allowing only an average evaluation of network performances during the analysis period.

However, this paper can be seen as preparatory for the implementation of a real-time incident management software for public transport networks. In fact, the outcomes of this model can be used as inputs of the Transit Link Transmission Model (Gentile, 2017), which is a fast-macroscopic model that performs a dynamic network loading from the results of a static assignment in combination with the available demand, supply and service features.

References

1. Beckman, M., McGuire, C., and Winsten, C. (1956) *Studies in the economics of transportation*. Yale University Press.
2. Bellman, R. (1958) On a routing problem. *Q Appl Math*, 16:87–90.
3. Cascetta, E. (2009) *Transportation Systems Analysis Models and Applications*. Springer.
4. De Cea, J. and Fernandez, E. (1993) Transit assignment for congested public transport system: An equilibrium model. *Transportation Science*, 27(2):133–147.
5. Dijkstra, E.W. (1959) A note on two problems in connexion with graphs. Numer Math, 1:269–271.
6. Dora, C. and Phillips, M. (2000) Transport, environment and health. *WHO Regional Publications, European Series*, 89.
7. Florian, M., Florian, D., and Constantin, I. (2009) A new look at projected gradient method for equilibrium assignment. Transportation Research Record: *Journal of the Transportation Research Board*, 2090:10–16.
8. Gentile, G. (2016) Solving a Dynamic User Equilibrium model based on Splitting Rates with Gradient Projection algorithms. *Transportation Research Part B* 92, 120-147
9. Gentile, G. (2017) Formulation of the transit link transmission model. *Transportation Research Procedia*, 27:889–896.
10. Gentile, G. and Noekel, K. (2016) *Modelling Public Transport Passenger Flows in the Era of Intelligent Transport System: COST Act ion TU1004* (TransITS). Springer.
11. Gentile, G., Florian, M., Hamdouch, Y., Cats, O., and Nuzzolo, A. (2016a) The theory of transit assignment: basic modelling frameworks. In Gentile, G. and Noekel, K., editors, *Modelling Public Transport Passenger Flows in the Era of Intelligent Transport System: COST Act ion TU1004* (TransITS), chapter 6, 287–386. Springer.
12. Gentile, G., Noekel, K., Schmocker, J.D., Trozzi, V., and Chandakass, E. (2016b) The theory of transit assignment: demand and supply phenomena. In Gentile, G. and Noekel, K., editors, *Modelling Public Transport Passenger Flows in the Era of Intelligent Transport System: COST Act ion TU1004* (TransITS), chapter 7, 387–439. Springer.
13. Kurauchi, F., Bell, M.G.H., and Schmocker, J.D. (2003) Capacity constrained transit assignment with common lines. *J Math Model Algorithms*, 24:309–327.
14. Oxford Economic Forecasting (2003) The economic effects of transport delays on the city of London. *Corporation of the City of London*, pages 26–27.
15. Rosen, J.B. (1960) The gradient projection method for nonlinear programming. Part I. linear constraints. *Journal of the Society for Industrial and Applied Mathematics*, 8:181–217.
16. Spiess, H. and Florian, M. (1989) Optimal Strategies: A new assignment model for transit networks. *Journal of Transportation Research*, 23B(2):83-102
17. WHO/Europe (2017) *Transport and health*. - http://www.euro.who.int/en/healthtopics/environment-and-health/, last visited Dec 23, 2017.