Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia - Social and Behavioral Sciences 54 (2012) 1145 – 1155

Procedia
Social and Behavioral Sciences

EWGT 2012

15th meeting of the EURO Working Group on Transportation

A review on transit assignment modelling approaches to congested networks: a new perspective

Qian Fu*, Ronghui Liu, Stephane Hess

Institute for Transport Studies (ITS), University of Leeds, Leeds LS2 9JT, United Kingdom

Abstract

This paper reviews a number of studies on both frequency- and schedule-based transit assignment models that have been proposed by far, wherein various behavioural assumptions on a wide range of aspects are embedded. With a reinvestigation on the relationships and homogeneity between different modelling approaches, it explores the representative veins of the models, and thereby extends a new perspective to the existing reviews under a historical context. Meanwhile, both advantages and disadvantages of these methods are presented. On the strength of the analyses and discussions of the state-of-the-art transit assignment models, further research directions are suggested.

Keywords: public transport; transit assignment models; congested transit network; path choice; travel behaviours

1. Introduction

With booming travel demand, congestion and overcrowding on transit network occurs frequently during rush hours. It has significant impact on regularity and reliability of public transport service, and influences passengers' travel behaviours. Travellers' perceptions differ in system performance and sensitivities to value of travel time, and hence different path choices are made. For decades, these issues have been long challenging and motivating transport planners and researchers to contrive transit assignment models, which (a) predict the passenger volumes choosing different services connecting any pairs of origin and destination (O-D); and (b) validate the effectiveness of operation schemes for transit systems. In that sense, the transit assignment problem has been well interrogated, modelled and implemented from two distinct standpoints: *frequency-based* (also termed headway-

* Corresponding author. Tel.: +44 (0)113-343-1790; fax: +44 (0)113-343-5334.
E-mail address: tsqf@leeds.ac.uk

 doi:10.1016/j.sbspro.2012.09.829

based or line-based) approach (e.g. [1][2][3][4][5][6]); and *schedule-based* (also termed timetable-based or run-based) approach (e.g. [7][8][9][10][11][12]). So far, there is a wealth of research gaining insights into both methods, wherein various assumptions on a wide range of aspects are embedded into different stages of the models' building process. Among the earlier surveys of transit modelling approaches are contributed by Bouzaïene-Ayari et al. [13][14]. According to Bouzaïene-Ayari et al. [13], the function that underlies the existing transit assignment models can be summarised in three aspects: (a) characteristics of the supply on transit networks and services; (b) information about the supply that passengers could have before and during their journeys; and (c) passengers' responses towards current situations given related travel information. Schmöcker [15] and Teklu [16] conduct reviews in more details of the frequency-based models. Nuzzolo, et al. [17] and Nuzzolo & Crisalli [18] concentrate on schedule-based transit assignment models and particularly elaborate the differences of the adaptability of schedule-based models to services with low frequency and high frequency. Furthermore, Nuzzolo & Crisalli [19] extend the formers to a wider scope involving multi-modal transportation networks of both transit and freight services. More recently, Liu et al. [20] inspects a great variety of studies on passengers' route choice behaviours with transit assignment approaches, ranging from conventional static models to dynamic ones given the effect of real-time information.

This paper reinvestigates the relationship and homogeneity of the behavioural assumptions among different transit assignment models. The fundamental premises of these models are expounded in Section 2 serving as the groundwork that sustains the subsequent sections. Section 3 presents the comparative edges and discussions of different assumptions embedded in relevant modelling approaches. In Section 4, it is explored that the veins of different models and a new perspective is thereby extended to the existing reviews under a historical context. Finally, on the strength of the discussions and analyses of the state-of-the-art transit assignment models, further research directions are proposed in Section 5 that concludes this paper.

2. The Backbone

2.1. Transit network, path, and paths

To reproduce passengers' travel behaviours, particularly to find out their choices of travel paths on a transit network, it is indispensable to reveal their decision-making processes at every decision-making point with onward journey segments in sequence. But different assumptions are likely to bring about different travel cost for each journey segment hence the total for a travel path.

A transit network can be represented under a graphic framework, and passenger flows are transmitted between decision nodes via different functional arcs, respectively. Consider that a complete journey in rail transit is segmented into successive parts, involving access to a stop (and buy a ticket if necessary), walking to a platform, waiting to board, travelling on-board; alighting (and/or interchange to another line if necessary) and egress. Each of these journey segments is associated with a travel time (or travel cost); and a passenger's journey cost could be generally regarded as the (weighted) sum of costs of all journey segments.

The transit network is simplified in many studies that only look at "stop-level" O-Ds, where the start stop is taken for the origin (centroid) with the end stop being the destination (centroid). In that case, a sub-network representing a transit stop becomes the focal issue (e.g. [14][21]). Passengers will choose one of attractive lines at each stop and travel to/through the next stop. An "attractive line" is commonly defined as a service that minimises a traveller's expected total journey time [18]. Different combinations of the attractive lines that the passenger takes at each of the successive stops result in different paths and thus a set of paths exist in a transit network, each of which links an O-D pair. Note that the passenger's choice of an attractive line is made at a transit stop, whereby the passenger flows split among the attractive lines. A more specific definition of attractive lines is given by De Cea & Fernández [22], which suggests that not all the available lines at a platform would be considered by a wait-to-board passenger, but only a portion of (i.e. a subset of) all lines that are available is perceived as rational. Those passengers would simply ignore the rest that can be obviously leading on "bad" path/paths. On top of this, the mostly accepted assumption for boarding is that, in uncongested situations, all passengers would board the firstly arriving line out of attractive ones.

2.2. Fundamental premises

The core that premises the outcome of transit assignment models is that a traveller always wishes to choose a cost-efficient path to complete his/her journey. For any of the passengers the journey cost is supposed to be minimised or optimal corresponding to the path he/she uses. Such cost may be dealt with a single path in earlier models (e.g. [23][24][25]), or a hyperpath/strategy, i.e. a probabilistic cost over a set of elementary paths that are considered simultaneously by the passenger (e.g. [2][3]). On the one hand, it could be simply regarded as the total travel time through the journey (i.e. the sum of observations or estimates for the travel time at every journey segment). On the other hand, in a synthetic manner, a generalised cost that takes into account not merely travel time but also other stochastic attributes and uncertainties up to the complexities of modelling perspectives of analysts [26][27], such as reliability of transit services, crowding, value of time and seat availability, as well as passengers' perceptions to such issues and so forth.

Provided that any additional information (e.g. remaining waiting time for a certain service) is not available, a passenger will always choose to board the firstly arriving run if alternative services are absent. When the common lines problem exist (e.g. [1][23][28]), whether to board the coming run of an attractive line, keep staying at the stop but waiting for the next run of the same line or another, or transfer to any of alternative services, all depends on what information would be involved with the passenger's decision-making process. In fact, passengers may not be able to know exactly the true journey cost of each alternative route but merely estimate and make trade-off decisions within their choice sets. For example, when travelling on board, a passenger needs decide whether to stay onboard or alight-to-transfer at the approaching stop, and how to continue at intermediate stops. These successive decisions lie with individual consideration in regard to the traveller's own perceived choice set, the service reliability and irregularity, and the availability of information (particularly the real-time information) the traveller can get prior to and in the course of their journeys. Several comparative scenarios of information effecting passengers' choice behaviours are discussed by Nökel and Wekeck [29][30]. They demonstrate that different assumptions made upon these issues would give rise to diverse models of route choice concerning both boarding and alighting occasions, and hence result in different splits of passenger volume into different lines.

Further, the waiting time for boarding a transit vehicle at a stop/platform is an indispensable factor that needs special treatment. For the frequency-based approach, headways between (or service frequencies of) lines/runs are fundamentally involved with passengers' wait for a transit vehicle's arrival. As the headway and service frequency are reciprocals, a passenger's average waiting time is supposed to be half of an average headway (or half of the inverse of an average service frequency). In this respect, different assumptions made on the headways or frequencies can deduce variable waiting times that the passenger will possibly spend at the stop/platform, which may then induce different path choices. However, dynamic attributes that are time-dependent cannot be taken into account explicitly, such as changing travel demand by different periods of a day and between days and passengers' arrival/departure times at stops, because a frequency-based transit network is represented within space domain only. In typical cases, it can hardly reveal the continuously accumulated loading burden of transit vehicles due to peaked congestion. For instance, passengers may fail to board under crowding situations, but keep waiting on the platform instead. Certainly, it aggravates the crowding on platform and hence affects the service that follows at the subsequent scheduled time interval. The increasingly intensive congestion may maintain in a certain period (e.g. morning peak hours).

In view of the scheduled-based models, every move of the entities (i.e. transit vehicles and passengers) is marked with a timestamp, such that those entities can be located, described, and differentiated from each other in both the temporal and the spatial dimension. Representation of the transit network is thereby adapted from line-based spatial-only graph (i.e. without time dimension), into a run-based spatio-temporal graph that can show individually serial runs as scheduled in timetables and more than competitive lines. Therefore, the characteristics of each service run can be inherently accommodated and modelled separately. Such run-based structure accordingly requires more capacity to be computationally tackled. Additionally, individual passengers, especially those frequent users of public transport, gain travelling experience of the transit system day by day, own different personal characteristics, and therefore have respective perceptions to making travel decisions on the choices of path and departure time, etc. Since an individual's arrival time at a stop/platform is considered as well as that of all runs, his/her waiting times are deterministic for reliably punctual services under uncongested situations, and time-variant trip demands and passenger loadings onto each of vehicles/runs can be shown in a more realistic manner through schedule-based approaches.

3. Behavioural Assumptions and Path Choice

From the above, service supply and travel demand is changing over time and interplays by the force of passengers' choice behaviours that in turn react to the interactions of the formers. This section goes deeper insights accounting for the related behavioural assumptions throughout the passengers' journeys. It compares horizontally different assumptions of choice behaviours, and various transit assignment modelling approaches are discussed. Note the hypothetical viewpoints might be slightly different from urban rail to bus transport, such that the term "platform" is distinguished from "stop". The terms stop and station are interchangeable.

3.1. Interactions between supply and demand

To look at the supply side of the transit network, it is emphasised on the attributes of transit network and line services, e.g. stops, alternative paths, service frequencies or timetables, and capacities. On the demand side, passenger flows distributed over all sections of the transit network tend to fluctuate from time to time, which generates different travel patterns by different periods of a day and days. Due to interactions between the supply and the passengers' choice behaviours that are in effect stage-by-stage throughout their own journeys, a construction of the entire travel process is often referring to the theory of strategy/hyperpath. Spiess and Florian [3] specify a strategy as a set of rules that enable a passenger to reach the destination. It corresponds to sequential decisions of choosing among attractive lines at every intermediate stop en-route to carry on with his/her journey. Note that, in this regard, a strategy is essentially a hyperpath that involves a set of elementary paths in light of the definition by Nguyen & Pallottino [2]. When multiple services exist, a set of strategies/hyperpaths can be enumerated and used by passengers based on their own considerations.

At practical O-D level, a passenger needs to make choices that involve stops, lines, and departure times and so forth, especially on schedule-based models. The travel cost of getting access to each station have thus to be taken into consideration in the formulation of a hyperpath. Confronted by any passenger is a joint choice, which is also marked by e.g. Nuzzolo et al. [31] and Nuzzolo et al. [32][33]. It reflects that a passenger may face a dilemma, that is for instance, (a) the access to alternative stops may be scattered around the origin or the destination and are of similar proximities; and (b) alternative transit lines can be chosen, which are together serving a stop or located at those scattered stops. When there is not accurate information being provided, the choice set of stops is largely correlated with that of lines, in particular for a less-experienced traveller. A transit map then becomes a major source that the passenger can use for reference to make path choices. It does have the utmost impact on his/her travel strategies [34]. Whereas, an experienced passenger, e.g. a commuter, has a choice set that is relatively fixed. If the predetermined path is no longer attractive a passenger may turn to another alternative path that hence newly becomes the optimal.

Most of the existing frequency-based models deal with the passenger assignment onto a hyperpath at stop-level O-Ds. Consider the transit network with high service frequency. The bulk of travellers may not concern too much that whether there would be a vehicle available to catch as soon as they arrive at a transit stop. A short wait would be acceptable for the majority of travellers. In contrast, when the targeting line serves less frequently, it would be necessary for those passengers to plan their departure times in order to minimise the waiting time for their targeted vehicles arriving at the platform, otherwise they will have to miss the targeted run. In addition, an individual passenger’s arrival at a stop/platform is commonly considered as independent of others under high-frequency service; and it does not correlate to any vehicle’s arrival, either. These two hypotheses embedded in many existing models determine the distribution of passengers’ arrival rate is uniformly random, which in turn underlies the assumption that passengers always choose to board the firstly arriving vehicle that belongs to attractive line set, given a Poisson process of transit vehicles’ arrival.

In frequency-based models that are irrespective of the common lines problem, the headway is obviously considered as the time interval between two consecutive transit vehicles that serve for the same line, with the mean being the average waiting time for that line. While the common lines problem is included, vehicles of different lines arrive at a platform alternately according to their respective headways. In this regard, the average inter-arrival time between runs is contracted due to the joint services, namely, a passenger's average waiting time for boarding (an attractive line) is dependent on a combined service frequency of all the attractive lines.

As mentioned in Section 2.2, the correlation between the arrival of vehicles and passengers yields different passengers’ waiting times, with either of the arrival rates drawn from specifically statistical distribution. Exponential distribution is the mostly used assumption prescribed for the headways. What’s more, Bouzaïene-Ayari et al. [14] argues it may not be appropriate in the case of reliable service regularity, as extremely irregular headways are not frequently encountered which however might be the case for exponential distribution. Erlang distribution is thus proposed and used for approximating the headways (e.g. [13][14]).

For the common lines problem, conventionally, the probability of boarding an attractive line is calculated as the proportion of its service frequency among all alternatives, which in turn splits the aggregated passenger flow. This implies that the more frequent a line service is, the higher probability that a vehicle of the transit line would be firstly arriving, and the greater chance it could obtain of being chosen. Note that any line available at a stop/platform is specifically associated with a path, and there could be a set of alternatives. Each alternative path can be assigned a probability of being chosen, even though illogical ones are never used that have zero probabilities. Another important issue mentioned in Section 1 is recalled that whether group-passenger choice behaviours or individual passenger choice behaviour is expected to be modelled or simulated. For the group passenger behaviours, it is assumed that, at each of reached stops, a group (or a certain type) of passengers is sharing an identical set of attractive lines. All the arriving passengers are classified into different groups each of which the member passengers will board the same line. While for the individual passenger behaviour, everyone considers his/her own path choice set that can be arguably assumed as independent of each other. Micro-simulation approaches are required to accommodate such personal features.

3.2. Capacity limitation, congestion and overcrowding issues

Passengers may experience congestion when they are travelling in carriages, waiting on the platforms and walking in the pedestrian channels that link platforms and station entries/exits. It occurs because the network supply is not able to serve for the excess travel demand during a given operational interval, which is typical of rush hours such as morning and evening peaks. In particular, when engineering works are scheduled or any incidents are encountered, service delays may be caused that hinders the system from releasing the surges of incoming passenger flows. Sometimes, overcrowding obstructs additional travel demand so that passengers’ travel strategies may alter. It also has impacts on in-vehicle passengers’ decisions of alighting as passengers may tend to concern more the possibility of being seated during their (long-distance) journeys. Some others may thus avoid the peak periods purposely with their choices of departure times being either earlier or later.

3.2.1. Vehicle capacity and boarding failures

Passengers' travel behaviours and path choices are enormously influenced by congestion and overcrowding, which affects the model estimates of passenger loadings onto each service run. In practice, service capacity associated with each journey segment has limited ability to manage on the flows of passenger volumes. Consequently the vehicle loadings that are either allocated into the vehicles or gathered in stops are supposed to be controlled below the capacity limits by assignment models.

For uncongested models, the intend-to-board passengers are all able to board one of their attractive line vehicles as the line capacity is not strictly limited. If the timetable is taken into account along with passengers' arrival time [7], an individual's choice of any run becomes deterministic and the common lines problem vanishes [35]. However, in congested cases, there are passengers failing to get on the vehicles.

Fail-to-board passengers can be generally classified into two groups of those (a) who fail to board due to limited standing space; and (b) who are not willing to board because of unavailability of seats. It corresponds to two situations. For one thing, the capacities of the carriages have been completely fulfilled and the wait-to-board passengers are unable to board. For another, a portion of all the wait-to-board passengers are unwilling to get involved into the crowded carriages hence avoid getting onboard even if space for boarding is still available. In addition to crowding-averse passengers, the rest who are seat-sensitive would prefer to be seated travelling. In such cases, passengers may not be in a position to board the first arriving vehicle.

De Cea & Fernández [22] considers the waiting time is monotonically increasing as the passenger volume increases and the volume-dependent waiting time is treated with congestion cost function that relates to the notion of effective frequency (e.g. [4][5]). An effective frequency is to characterise an attractive line or common lines. That is, if the probability that a passenger encounters a full vehicle goes up, the effective frequency of the relevant service decreases, and hence the waiting time for that vehicle/line to arrive becomes longer. However, the congestion cost function does not actually restrain line or vehicle capacities from being overloaded by excessive travel demand. In another way, the method that specifies explicit constraints is employed to prohibit strictly the superimposed passenger flows being assigned onto any path sections with limited capacities (e.g. [36][37][38][39][40][41][42]). In practice, the capacity of a transit vehicle as well as a stop/platform is limited but can only accommodate certain amount of passengers demand for boarding, alighting and interchange. It is important to understand the vehicles' loadings (i.e. numbers of onboard passengers) when they are arriving at a platform; and through the choice processes of all the wait-to-board passengers, the updated vehicles loadings when the vehicles are departing from the platform. This information is a prerequisite of adapting the flow-dependent cost function for the journey segments of wait and onboard travelling, through which a traveller's discomfort with the journey is supposed to be going up with the increasing passenger flow.

Besides, the probability of failing-to-board that affects the search for the shortest hyperpath is specially discussed by e.g. Kurauchi et al. [43], Schmöcker et al. [44] and Schmöcker & Bell [45]. The choice set of lines considered by passengers who fail to board may change in different time intervals, and it depends only on the current condition, which is known as Markov property and also presented by Teklu et al. [46]. Fail-to-board passengers who keep waiting on the same stop obey with the Markov properties; and hence, whether or not they are able to board the currently arriving vehicle is not related to where they started their journey or how long they have been waiting. The boarding and alighting demands at the current stop are necessary, which requires the knowledge of the traffic volumes at the upstream stops each associates with a timestamp. Consequently, the waiting time at a given stop/platform depends on the variations of traffic volume over time or time intervals, and the current traffic volumes both on-board and at the stop would lengthen the waiting time practically.

3.2.2. Seat availabilities

Recently, the issue of seat availability has been taken into account that influences passengers' travel strategies under less congested circumstances; whereas, this is not the case when the network is suffering from high congestion during periods of peak demand. Because in highly congested conditions whether a passenger could be seated onboard would not be the main concern. Instead, whether there is a chance for passengers to board would be valued given that the onboard crowding does not outstrip passengers' tolerance limits to congestion while the vehicle still has capacity of added boarding demand. A transit vehicle that is less congested may be more attractive to some passengers, even though it may cost a longer travel time compared to the alternatives [47]. From the perspective of value of time, some passengers who fail to board shall have a longer perceived travel time if the chosen line is more crowded on-board. If they transfer to another line (or keep waiting for the following line) with a longer in-train time but less onboard crowding, their perceived travel time would be equal to or even shorter than the pre-planned, although an extra elapsed waiting time in vain is counted.

Another problem of blending "seat congestion" into the transit assignment models is the specification of who has the priority of being seated among passengers, particularly the seat-sensitive ones. A set of priority rules are proposed in an empirical view by numerous studies (e.g. [47][48][49]) and the general principles are as follows:

- Passengers already sitting onboard will maintain the seats until they alight.
- Passengers who are standing on-board and keep on travelling with the same line are prioritised to hunt for seats over the wait-to-board passengers on the platform; and the former are assumed to take any seats that are freed before the latter board.
- The standing-on-board passengers have the same chance of getting any vacant seats.
- The passengers waiting on the platform board a vehicle on the First-Come-First-Serve (FCFS) basis.

It is specified by Schmöcker et al. [50] that all passengers waiting on the platform compete coequally for those available seats rather than complying with FCFS. It implies that one every wait-to-board passenger has the same chance of boarding as well. This is more reasonable for urban rail cases in less- or un-congested circumstances, because late-arriving passengers can select to wait by any carriage doors.

Suppose that a vehicle is approaching or has already arrived at a stop. A standing-on-board passenger will be looking for a seat that would become available. On the one hand, when the carriage is crowded, the seated passenger would move closer to the door, and seats become available for any standing one and are reallocated before potential travellers board the same carriage. On the other hand, travellers who have been waiting on the platform will start boarding as soon as the passengers onboard who intend to alight are all cleared. Those intend-to-board passengers can estimate how much boarding capacity remains on the vehicle and consider whether there will be a chance of being seated. They may also estimate the probability of obtaining vacant seats at the subsequent stop(s), and thereby path choices will be made conditional on whether the passenger is going to board or keep waiting at the current stop.

Nevertheless, the degree of passengers' incentives of picking up vacant seats differs; and it is hardly quantified, which generally involves two stages: before boarding (the first) vehicle and during standing in-vehicle. Before a passenger board, the key influence factors involve the total journey distance and the occupancy of the arriving vehicle's capacity. The latter can be measured by the ratio of current passenger volume to the vehicle capacity. In another aspect, suppose the passenger has been standing and travelling on-board. The elapsed time he/she has already spent in standing and its interplay with the remaining journey length then become the main influence factor. Such stimulus can be therefore scaled in terms of the Nested Logit structure under which the mentioned attributes correlate. However, the existing models assume that standing passengers have an equal probability of being seated if a vacant seat is available, and so do those newly boarding passengers with lower priority.

The third issue concerns the different travel discomforts experienced by standing and sitting passengers (e.g. [47][48][49][50][51]). It leads to different formulations of the in-vehicle travel costs, respectively, and hence adjusted different travel strategies. While being crowded on-board, the degree of discomfort is much more highly valued by passengers who are standing than those who are seated; and the latter are assumed scarcely influenced but having similar level of discomfort as they travel under less- or un-congested environment.

4. Representative Veins of Models

In the case of passengers failing to board, Schmöcker & Bell [52] present a frequency-based transit assignment model taking account of passengers' boarding failure incurred by the insufficient room in vehicle, which thereupon causes surplus waiting time of the wait-to-board passengers on the platform. This undesired delay certainly affects the passengers' choices of paths, regarding the perceived cost of the whole journey. Built on properties of Markov processes is introduced "notational bins" considered as an absorbing states to gather the exceeding passenger volumes that fail to board the vehicle due to the shortage of boarding space.

Based on [52], Kurauchi et al. [43] considers not only transit line capacities, but also the common lines problem that was not implemented in the former. The common lines problem is explained by the virtue of the network's graphic representation of "stop nodes", and the excess demands of fail-to-board passengers are sent directly to their respective destinations by "failure arcs". At a "fail-to-board" node, a proportion of wait-to-board passengers can successfully get on a train while the rest are left on the platform. That fail-to-board demand is therefore associated to a fail-to-board probability. The split probabilities at the "alighting nodes" are not discussed in this study because, normally, passengers would not alight and then board the same line. However, interchange was not considered in this model. Passengers may possibly alight to transfer to another line, in which case, extra time for interchange on "transfer arcs" may be required. While a train is dwelling on a platform, the cost on "stopping arcs" is not equivalent to dwell time that includes passengers' alighting time and boarding time. As the alighting passengers have priority of occupying the doors, the process of boarding starts as soon as that of alighting is cleared. Therefore, the dwell time starting from the moment that a train stops at a platform consists of alighting time, boarding time and time of clearance confirmation before the train departs. It is counted in each passenger's waiting time rather than being considered separately. The common lines problem for urban rail service can be reduced or avoided when the platform is line-specific, in which case those competitive lines are all serving the same station but will stop at different platforms; and passengers waiting on each platform have already decided which line to take. Then the problem goes to a line choice model that determines to which platform a passenger would head. The probabilities of heading to platforms are vital for the first round of passenger splits in transit assignment process. Kurauchi et al. [43], Bell & Schmöcker [53], Schmöcker et al. [44], and Schmöcker & Bell [45] contribute progressively in the assignment method and solution algorithm for the issue of failure-to-board, among which Schmöcker et al. [44] and Schmöcker & Bell [45] account for the time-dependent boarding demand by extending the former models. They consider different time intervals that show the dynamic nature of the modelling approach. That is, those fail-to-board passengers within the current time interval (i.e. the absorbed demand in [43]) still keep waiting on the platform and mingle with the passengers who wish to board in the subsequent time interval (i.e. updated boarding demand). However, it is supposed that the probability of failing to board maintain constant given a time interval. Schmöcker et al. [50] injects the effect of seat availability into the predecessor models mentioned above by introducing a fail-to-sit probability. Thus, standing and sitting passengers are naturally distinguished in the network representation and model specification. The in-vehicle travel cost of standing passengers is subject to a standing penalty when considered in journey cost of the hyperpath. Nevertheless, the line/vehicle capacity constraint is not guaranteed in the model.

Nuzzolo and Russo [54] firstly presents the "space-time graph" to model low-frequency transit services for which dynamic approaches were fit only in according to common practice. But it is argued that schedule-based modelling approach can be a great fit for high frequency cases as well [31]. Established on the proposed diachronic network representation, the latter adapt more complex details into a "doubly" assignment model that involves both within-day and day-to-day dynamics for transit network with high frequency of services. In reality, the actual arrival/departure time of a run may not be consistent with the schedule but vary from time to time. It is then considered as random with the published scheduling time being the mean. Nuzzolo et al. [11] and Nuzzolo et al. [32][33] further take into account the capacity issues at run-based level in schedule-based modelling framework. Nuzzolo et al. [33] presents the onboard-travelling, boarding and alighting demands of each service run at each stop where runs are individualised in terms of their scheduling arrival/departure times. Thus, a passenger's earlier or later arrival at the stop than that of a run is explicitly related to which runs he/she would consider and which one is most likely to be chosen eventually.

In another regard, journey time variability is often employed to scale to the transit service reliability that also exerts significant influences on passengers' trip decisions upon travel strategies, as well as their perceptions to the system performance. Travellers are always trying to make trade-off choices in the context of highly uncertain transit service. Note that not only is journey time variability considered as in effect for in-vehicle travel and wait, however, it is also observed in walk and any other stages of a complete journey (e.g. [26][27][55]). It is pointed out by Szeto et al. [26] that the effect of the variations of generalised journey time is not given much treatment into the decisions of route choices in the existing transit assignment models.

When a transit vehicle/run breaks down in transit or terminate at certain station due to system fault, passengers could either choose to keep waiting onboard, or shift to any alternative lines, or even egress and go for any other modes. However, if no alternative services are available, then the passenger will need to wait until the fault is cleared or for the next coming train. In high frequency service, seconds of delay would result in a series of delays of the following runs, which in turn cause reallocation of passenger flows.

5. Conclusion

This paper illustrates different perspectives of looking at congested transit assignment problems. A number of studies on both frequency- and schedule-based transit assignment models have been reinvestigated, especially with respect to issues of congestion and overcrowding. Comparisons of travellers' choice behaviours and the modelling approaches to transit assignment are presented and discussed in horizontal and longitudinal views. Rooted in the above-mentioned explanatory frameworks of the shared features, distinct emphases are placed on different layers of travelling processes to interpret passengers' behaviours. Different veins of transit modelling approaches are thereby probed as well.

At the within-day level, network status shifts in the light of distinct periods, and transit models shall be required to integrate corresponding perspectives and approaches to replicate the scenarios, respectively, from uncongested, less congested, to fully congested environments along with a process of congestion release.

Furthermore, referring to Schmöcker et al. [50], it is important to combine the "fail-to-sit" and "fail-to-board" probabilities into one modelling framework considering simultaneously the capacities of each run and vehicle's seats.

Finally, assumptions of the process that a passenger determines individually his/her perceived path choice set is one of the issues that need to be addressed in future work, especially pertaining to micro-simulations. On the other hand, from the microscopic perspectives, how the passenger chooses to board a line out of his/her choice set needs to be further discussed, given the effect of seat availability. It will be thus necessary to distinguish those wait-to-board, intend-to-board, decline-to-board and seat-sensitive passengers, to accommodate the fact of first refusal concerning e.g. unwilling-to-stand and unwilling-to-sit cases.

Acknowledgements

The authors appreciate funding support from China Scholarship Council (CSC) and University of Leeds, and the reviewers' comments.

References

[1] Chriqui, C., & Robillard, P. (1975). Common bus lines. *Transportation Science, 9*, 115-21.
[2] Nguyen, S., & Pallottino, S. (1988). Equilibrium traffic assignment for large scale transit networks. *European Journal of Operational Research, 37*, 176-86.
[3] Spiess, H., & Florian, M. (1989). Optimal strategies: a new assignment model for transit networks. *Transportation Research Part B: Methodological, 23*, 83-102.
[4] Wu, J. H., Florian, M., & Marcotte, P. (1994). Transit equilibrium assignment: a model and solution algorithms. *Transportation Science, 28*, 193-203.
[5] Cominetti, R., & Correa, J. (2001). Common-lines and passenger assignment in congested transit networks. *Transportation Science, 35*, 250-67.
[6] Cepeda, M., Cominetti, R., & Florian, M. (2006). A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria. *Transportation Research Part B: Methodological, 40*, 437-59.
[7] Tong, C. O. (1986). *A schedule-based transit network model.* Unpublished PhD, Monash University.
[8] Hickman, M. D., & Bernstein, D. H. (1997). Transit service and path choice models in stochastic and time-dependent networks. *Transportation Science, 31*, 129-46.
[9] Florian, M. (1999). Deterministic time table transit assignment. In *First Asian EMME/2 Users Group Meeting*. Shanghai, P.R. China.
[10] Tong, C. O., & Wong, S. C. (1999). A stochastic transit assignment model using a dynamic schedule-based network. *Transportation Research Part B: Methodological, 33*, 107-21.
[11] Nuzzolo, A., Russo, F., & Crisalli, U. (2001). A doubly dynamic schedule-based assignment model for transit networks. *Transportation Science, 35*, 268-85.
[12] Poon, M. H., Wong, S. C., & Tong, C. O. (2004). A dynamic schedule-based model for congested transit networks. *Transportation Research Part B: Methodological, 38*, 343-68.
[13] Bouzaïene-Ayari, B., Gendreau, M., & Nguyen, S. (1998). Passenger assignment in congested transit networks: a historical perspective. *Equilibrium and advanced transportation modelling*, 47-71.
[14] Bouzaïene-Ayari, B., Gendreau, M., & Nguyen, S. (2001). Modeling bus stops in transit networks: a survey and new formulations. *Transportation Science, 35*, 304-21.
[15] Schmöcker, J.-D. (2006). *Dynamic capacity constrained transit assignment.* Imperial College London, London, UK.
[16] Teklu, F. (2008). *A Markov process model for capacity-constrained transit assignment.* University of Leeds, Leeds.
[17] Nuzzolo, A., Russo, F., & Crisalli, U. (2003). *Transit network modelling. The schedule-based dynamic approach*. Milano: FrancoAngeli.
[18] Nuzzolo, A., & Crisalli, U. (2004). The schedule-based approach in dynamic transit modeling: a general overview. In N. H. M. Wilson & A. Nuzzolo (Eds.), *Schedule-Based Dynamic Transit Modeling: theory and applications* (1 ed., pp. 1-24): Springer.
[19] Nuzzolo, A., & Crisalli, U. (2009). The schedule-based modeling of transportation systems: recent developments. In *Schedule-Based Modeling of Transportation Networks* (pp. 1-26): Springer US.
[20] Liu, Y., Bunker, J., & Ferreira, L. (2010). Transit users' route-choice modelling in transit assignment: a review. *Transport Reviews, 30*, 753-69.
[21] Billi, C., Gentile, G., Nguyen, S., & Pallottino, S. (2004). Rethinking the wait model at transit stops. In *The 5th Triennial Symposium on Transportation Analysis*. Le Gosier, Guadeloupe.
[22] De Cea, J., & Fernández, E. (1993). Transit assignment for congested public transport systems: an equilibrium model. *Transportation Science, 27*, 133-47.
[23] Dial, R. B. (1967). Transit pathfinder algorithm. *205*, 67-85.
[24] Fearnside, K., & Draper, D. P. (1971). Public transport assignment - A new approach. *Traffic Engineering and Control, 13*, 298-9.
[25] Le Clercq, F. (1972). A public transport assignment method. *Traffic Engineering and Control, 14*, pp. 91-6.
[26] Szeto, W. Y., Solayappan, M., & Jiang, Y. (2011). Reliability-based transit assignment for congested stochastic transit networks. *Computer-Aided Civil and Infrastructure Engineering, 26*, 311-26.
[27] Szeto, W. Y., Jiang, Y., Wong, K. I., & Solayappan, M. (In Press). Reliability-based stochastic transit assignment with capacity constraints: formulation and solution method. *Transportation Research Part C.*
[28] Lampkin, W., & Saalmans, P. D. (1967). The design of routes, service frequencies, and schedules for a municipal bus undertaking: a case study. *OR, 18*, pp. 375-97.
[29] Nökel, K., & Wekeck, S. (2007). Choice models in frequency-based transit assignment. In *European Transport Conference 2007.* Leeuwenhorst Conference Centre, The Netherlands.
[30] Nökel, K., & Wekeck, S. (2009). Boarding and alighting in frequency-based transit assignment. *Transportation Research Record: Journal of the Transportation Research Board, 2111*, 60-7.
[31] Nuzzolo, A., Crisalli, U., & Russo, F. (1999). A doubly dynamic assignment model for congested urban transit networks. In *European Transport Conference 1999* (Vol. 434). Cambridge.
[32] Nuzzolo, A., Crisalli, U., & Rosati, L. (2007). Congested transit networks: a schedule-based dynamic assignment model with explicit vehicle capacity constraints. In *Proceedings of European Transport Conference*. Leeuwenhorst, The Netherlands.

[33] Nuzzolo, A., Crisalli, U., & Rosati, L. (2012). A schedule-based assignment model with explicit capacity constraints for congested transit networks. *Transportation Research Part C, 20*, 16-33.
[34] Guo, Z. (2011). Mind the map! The impact of transit maps on path choice in public transit. *Transportation Research Part A, 45*, 625-39.
[35] Wong, S. C., & Tong, C. O. (1998). Estimation of time-dependent origin-destination matrices for transit networks. *Transportation Research Part B: Methodological, 32*, 35-48.
[36] Lam, W. H. K., Gao, Z. Y., Chan, K. S., & Yang, H. (1999). A stochastic user equilibrium assignment model for congested transit networks. *Transportation Research Part B: Methodological, 33*, 351-68.
[37] Nguyen, S., Pallottino, S., & Malucelli, F. (2001). A modeling framework for the passenger assignment on a transport network with timetables. *Transportation Science, 35*, 238-49.
[38] Lam, W. H. K., Zhou, J., & Sheng, Z.-h. (2002). A capacity restraint transit assignment with elastic line frequency. *Transportation Research Part B: Methodological, 36*, 919-38.
[39] Yin, Y., Lam, W. H. K., & Miller, M. A. (2004). A simulation-based reliability assessment approach for congested transit network. *Journal of Advanced Transportation, 38*, 27-44.
[40] Teklu, F. (2008). A stochastic process approach for frequency-based transit assignment with strict capacity constraints. *Networks and Spatial Economics, 8*, 225-40.
[41] Hamdouch, Y., Marcotte, P., & Nguyen, S. (2004). Capacitated transit assignment with loading priorities. *Mathematical Programming, 101*, 205-30.
[42] Hamdouch, Y., & Lawphongpanich, S. (2008). Schedule-based transit assignment model with travel strategies and capacity constraints. *Transportation Research Part B: Methodological, 42*, 663-84.
[43] Kurauchi, F., Bell, M. G. H., & Schmöcker, J.-D. (2003). Capacity constrained transit assignment with common lines. *Journal of Mathematical Modelling and Algorithms, 2*, 309-27.
[44] Schmöcker, J.-D., Bell, M. G. H., & Kurauchi, F. (2008). A quasi-dynamic capacity constrained frequency-based transit assignment model. *Transportation Research Part B, 42*, 925-45.
[45] Schmöcker, J.-D., & Bell, M. G. H. (2009). The build-up of capacity problems during the peak hour. In *Schedule-Based Modeling of Transportation Networks: Theory and Applications* (pp. 217-39): Springer US.
[46] Teklu, F., Watling, D., & Connors, R. D. (2007). A Markov process model for capacity constrained transit assignment. In R. In Allsop, M. G. H. Bell & B. G. Heydecker (Eds.), *Transportation and Traffic Theory*. London: Elsevier.
[47] Leurent, F. (2010). On seat capacity in traffic assignment to a transit network. *Journal of Advanced Transportation*.
[48] Sumalee, A., Tan, Z., & Lam, W. H. K. (2009). Dynamic stochastic transit assignment with explicit seat allocation model. *Transportation Research Part B, 43*, 895-912.
[49] Hamdouch, Y., Ho, H. W., Sumalee, A., & Wang, G. (2011). Schedule-based transit assignment model with vehicle capacity and seat availability. *Transportation Research Part B, 45*, 1805-30.
[50] Schmöcker, J.-D., Fonzone, A., Shimamoto, H., Kurauchi, F., & Bell, M. G. H. (2011). Frequency-based transit assignment considering seat capacities. *Transportation Research Part B, 45*, 392-408.
[51] Tian, Q., Huang, H.-J., & Yang, H. (2007). Commuting Equilibria on a mass transit system with capacity constraints. In *Proceedings of the 17th International Symposium on Transportation and Traffic Theory (ISTTT)* (pp. 360-83): Elsevier: Oxford.
[52] Schmöcker, J.-D., & Bell, M. G. H. (2002). An application of congested transit network loading with the Markov chain approach. In (Vol. Handling Uncertainty in the Analysis of Traffic and Transportation Systems). Bari, Italy: Proceedings of the 9th Mini-EURO Conference.
[53] Bell, M. G. H., & Schmöcker, J.-D. (2004). A solution to the transit assignment problem. In N. H. M. Wilson & A. Nuzzolo (Eds.), *Schedule-Based Dynamic Transit Modeling* (pp. 263-79): Kluwer Academic Publishers.
[54] Nuzzolo, A., & Russo, F. (1994). An equilibrium assignment model for intercity transit network networks. In *Triennial Symposium on Transportation Analysis (TRISTAN) II*. Capri, Italy.
[55] Szeto, W. Y., & Solayappan, M. (2009). A variational inequality formulation for the reliability-based stochastic user equilibrium problem for congested transit networks. In *Proceedings of the Eastern Asia Society for Transportation Studies* (Vol. 7, pp. 87-102).