Public Transp (2010) 2: 69–86
DOI 10.1007/s12469-010-0027-4

ORIGINAL PAPER

# The multi-class schedule-based transit assignment model under network uncertainties

**Yuqing Zhang · William H.K. Lam · Agachai Sumalee · Hong K. Lo · C.O. Tong**

Published online: 6 August 2010


**Abstract** Demand and supply uncertainties at schedule-based transit network levels strongly impact different passengers' travel behavior. In this paper, a new multi-class user reliability-based dynamic transit assignment model is presented. Passengers differ in their heterogeneous risk-taking attitudes towards the random travel cost. The stochastic characteristics of the main travel cost components (in-vehicle travel time, waiting time, and early or late penalty) are demonstrated by specifying the demand and supply uncertainties and their interactions. Passenger route and departure time choice is determined by each passenger's respective reliability requirements. Vehicle capacity constraint for random passenger demand is handled by an in-vehicle congestion parameter. The proposed model is formulated as a fixed-point problem, and solved by a heuristic MSA-type algorithm. The numerical result shows that the risk-taking attitude will impact greatly on passengers' travel mode and departure time choices, as well as their money and time costs. This model is also capable of generating transit service attributes such as the stochastic vehicle dwelling time and the deviated timetable.

**Keywords** Reliability-based stochastic user equilibrium · Schedule-based transit assignment · Multi-class · Demand uncertainty · Supply uncertainty · Capacity constraint

Y. Zhang (✉) · W.H.K. Lam · A. Sumalee
Department of Civil and Structural Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China
e-mail: 07902133r@polyu.edu.hk

H.K. Lo
Department of Civil Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

C.O. Tong
Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, China

**JEL Classification** R410

## 1 Introduction

Transit assignment models are fundamental tools to enable transit system evaluation, control, and planning. Modeling techniques for transit assignment problems described in the literature are largely categorized as frequency-based (De Cea and Fernandez 1993; Cominetti and Correa 2001; Schmocher et al. 2008) and schedule-based (Wilson and Nuzzolo 2004; Poon et al. 2004; Hamdouch and Lawphongpanich 2008). These two modeling methods yet serve different planning purposes. The former aims at long-term planning such as land use and transport development projects, while the latter is better suited to short-term transit operations and service planning such as transit timetabling and vehicle scheduling.

Transit assignment models have recently emphasized the influence of uncertainties in frequency-based frameworks. The vertex failure in transit networks has been studied by Bell et al. (2002). A particular type of vertex failure (failure to board a full service) has been investigated by the absorbing Markov chain model. The notion of failure-to-board has been further applied in assignment models with common line problems (Kurauchi et al. 2003). Yang and Lam (2006) proposed a probit-type reliability-based transit assignment model for congested networks with unreliable transit services. Stochastic passenger in-vehicle travel time, impacted by vehicle on-road running time uncertainties was considered. Szeto et al. (2009) further considered the stochastic passenger waiting time and the stochastic capacity from the perspective of a line rather than a run. Stochastic passenger waiting time was due to random passenger arrival and stochastic distributed line headways, which has also been discussed by Spiess and Florian (1989). Stochastic vehicle capacity also stems from headway variation of the line.

The above frequency-based models can be used to study aggregated stochastic effects of a specific transit line from the static perspectives. However, uncertainties exist in both vehicle running and dwelling process in line operation. The influence of uncertainties is also different for each run. The schedule-based model provides a means to investigate uncertainties within the vehicle operation process. Nuzzolo et al. (2001) have investigated the dynamic transits system of regular and irregular services. The road congestion uncertainty resulting irregular service is merely defined exogenously. Teklu et al. (2007) studied the day-to-day passengers learning processes regarding stochasticity in transit networks by a micro-simulation-based approach. The variance of passenger perceived cost is given by an equation of line frequency and passenger in-vehicle travel time, but without justification. These models represent passenger perceived travel cost and network configuration under network uncertainties from dynamic perspectives, however they do not cover the evolvement and interaction of uncertainties and impacts of uncertainties on passenger travel behavior.

Sources of uncertainties can be divided into the supply side (the road conditions, weather, and incidents) and the demand side (within-day or day-to-day variation of passenger demand). As shown in Fig. 1, passengers boarding time variation and vehicle on-road running time irregularity are regarded as exogenous uncertainties. These

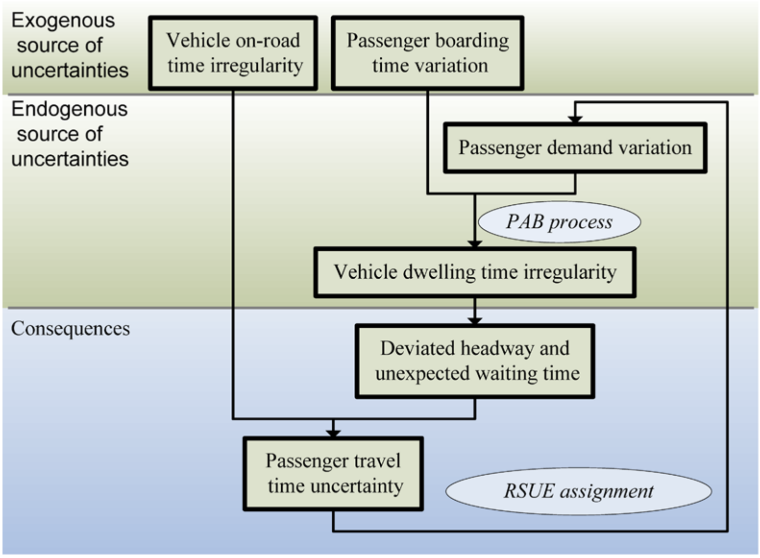



**Fig. 1** Sources, evolvements and interactions of transit network uncertainties

factors are independent of passengers' route and departure time choices, but nonetheless affect such choices. Hence the stochastic passenger demand is endogenously decided by passengers' reaction. Vehicle dwell time uncertainty is derived from the stochastic passenger boarding process, which is determined by the random passenger boarding time (exogenously) and the number of boarding passengers (endogenously). A stochastic passenger arriving and boarding (PAB) process is further described in Sect. 2 (transit service modeling).

Passengers may have different risk perception when facing above network uncertainties in their travels, as they may value travel time reliability differently, depending on their income levels and trip purposes (Noland and Polak 2002; Lam et al. 2008). Thus, passenger travel choice, such as mode, line, and departure time, varies in accordance with the heterogeneous risk-taking attitudes. In this study, multi-class users' different attitudes toward stochastic travel cost are considered as (1) risk-prone, (2) risk-neutral and (3) risk-averse, and modeled by a step function with anticipated possibility of on-time arrival. The stochasticity of travel cost is captured by using the effective travel time (ETT), allowing for a safety margin additional to the average travel time, so as to ensure the actual travel time remains within the time budget (Chen et al. 2002; Lo et al. 2006).

The stochastic travel cost consists of stochastic passenger in-vehicle travel time (composed of vehicle dwelling and running time), waiting time, early or late penalty, and out-of-pocket fares. Passenger stochastic waiting time can be the waiting time for the first arriving vehicle, or vehicles arriving thereafter due to the passenger's failure to board. Congestion and vehicle capacity constraints are the main reasons for the latter (overload delay). Under demand uncertainties, the deterministic physical vehicle constraint is adjusted by imposing a performance parameter to constrain the stochastic in-vehicle passengers.

This paper is organized as follows. Section 2 introduces the formulation of the proposed model. It includes the presentation of schedule-based stochastic transit network, the modeling of stochastic demand, the derivation of stochastic passenger travel cost, and the formulation of the reliability-based stochastic user equilibrium (RSUE) model. Section 3 demonstrates the dynamic transit network loading and the feasibility of a heuristic solution algorithm. In Sect. 4, a numerical example based on the transit network from the Kowloon area to Hong Kong International Airport is carried out to illustrate the application of the model, solution algorithm, and some important insights.

## 2 Model formulation

### 2.1 The schedule-based transit network representation

The schedule-based transit network is illustrated in the diachronic graph (Nuzzolo et al. 2003). Passenger movement and transit line running attributes (such as the schedule coordination and the vehicle encountering or overtaking phenomenon) can be represented in the graph by the time and space illustration of vehicle trajectories. Given a transit network $\Omega(I, J, L)$, the stochastic arrival time and departure time of vehicle $V_{i,j,l}$, the $i$th transit vehicle of line $l$ at stop $j$, are described as $T^a_{i,j,l}$ and $T^d_{i,j,l}$. The $i$th vehicle and $(i+1)$th vehicle of the same line may meet on road or stop due to the stochastic vehicle dwell time and on-road running time.

### 2.2 Modeling transit demand

The modeling time horizon $[0, T]$ has been divided into a vector of time intervals $\tau_i \in (\ldots, \tau_{i-1}, \tau_i, \tau_{i+1}, \ldots)$. It is assumed that the number of passengers arriving (including passengers just starting the trip and passengers transferring from other lines) at each time interval $\tau_i$ for each line $l$ is a inhomogeneous Poisson process $\{Q_l(\tau), \tau \geq \mathbf{0}\}$. When the number of time intervals is large enough, according to the central limit theorem, the cumulative passenger arrivals can be approximated by Normal distribution. Similarly, the cumulative passenger departures can be assumed to follow Normal distribution.

The passenger origin and destination (OD) demand should be equal to the total number of passenger arrivals minus transfers. Denote $D^{rs}$ as the passenger OD demand of the OD pair $r$ (origin) and $s$ (destination) over the investigating period, $D_l$ as the cumulative passenger arrivals at line $l$ and $T_l$ as cumulative passenger transfers at each line before reaching the destination. Assuming the independence of passenger arrivals and transfers, the OD demand can be represented as:

$$\sum_{r,s} D^{rs} = \sum_{l} (D_l - T_l), \tag{1}$$

which also follows the Normal distribution by the independent and identical distributed (IID) property.

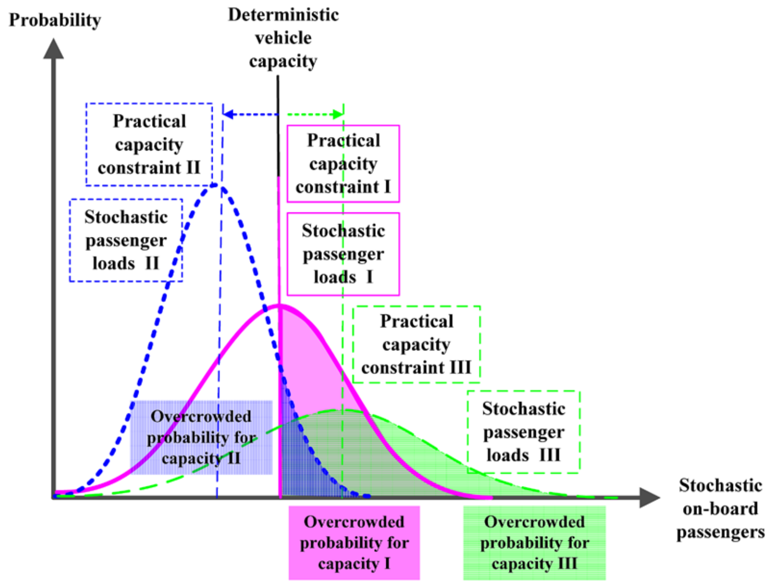



**Fig. 2** Stochastic passenger load under different capacity constraint

Wirasinghe (2003) revealed that bus load status can vary from underutilized to overloaded with respect to different dispatch time and elapsed travel time. In this stochastic network, where the possibilities of underutilizing and overloading are considered aggregately by passengers on-board, a pre-assumed overload parameter is used to represent the effect of vehicle capacity constraint. When the parameter is assumed to be 0.5, the designed capacity is used to bind the practical boarding passengers, but only indicates that in half situation, the vehicle is not overloaded (practical capacity constraint I in Fig. 2). Other values of the overload parameter could be calibrated to represent the congested or uncongested network (practical capacity constraint II and III in Fig. 2).

To specify the vehicle capacity constraint problem under the condition of demand uncertainty, the parameter $\omega$ is introduced to represent the probability of overloading. The equation then implies that the probability of passengers loaded in the $j$th run of line $l$ ($Q_l^j$) exceeds the vehicle capacity $cap_l$ by $\omega$:

$$P\{Q_l^j \geq cap_l\} = \omega. \tag{2}$$

The difference between on-board passengers and vehicle capacity is $\Delta CAP_l^j = Q_l^j - cap_l$. $cap_l$ is a constant and $Q_l^j$ follows Normal distribution, so the mean and standard deviation thereby can be written as: $E(\Delta CAP_l^j) = E(Q_l^j) - cap_l$ and $sd(\Delta CAP_l^j) = sd(Q_l^j)$. Standardizing this variable with the given confidence interval, the probability of overloading is:

$$P\{\Delta CAP_l^j \geq 0\} = 1 - \Phi\left(\frac{0 - E(Q_l^j) + cap_l}{sd(Q_l^j)}\right) = \omega. \tag{3}$$

Be reminded that the mean and variance are equally the square of the standard deviation of the number of passengers loaded under the assumption of Poisson distribution. Hence the standard deviation is substituted by the square root of the mean:

$$E(Q_l^j) + \sqrt{E(Q_l^j)} \cdot \Phi^{-1}(1 - \omega) - cap_l = 0. \tag{4}$$

The unique value of $E(Q_l^j)$ can be found by solving the square root equation, which reveals the practical capacity constraint in the stochastic network loading.

### 2.3 Modeling transit services

Passenger boarding and alighting varies in relation to largely the station design, door design, and fare collection method. For example, passengers who take buses usually use different doors for boarding and alighting while passengers who take trains may board and alight simultaneously in bulk and via the same doorway. Models have been developed and surveys conducted, focusing on the stochastic bus dwell time (Powell and Sheffi 1983; Adamski 1992; Lam et al. 1998; and Hickman 2001), but these results have not been applied to transit assignment models to estimate stochastic vehicle travel time.

Transit assignment models, considering vehicle dwell time owing to passenger service time (boarding and alighting time), were studied by Larrain and Muñoz (2008). The relationship between vehicle dwell time and the number of passengers boarding was regarded as deterministic and linear. However, the dwelling of transit vehicles is a stochastic process involving passengers arriving, queuing, and boarding (Zhang et al. 2009). Lam et al. (1998) investigated the train dwell time at several main stations in Hong Kong, and found that the train dwell time followed normal distribution. The alighting time, noticeably carries greater weight than the boarding time in the dwell time generalized equation.

The bus dwell time and train dwell time have considerably different mechanisms. On the outward trip at peak travel time, bus boarding time has a greater time weighting than the alighting time and, vice versa for the inward trip. It is also common for a bus to wait for a passenger rushing, belatedly to join the bus. As a consequence, the passenger arriving and boarding (PAB) process should only be modeled for outward bus trips; otherwise both the service time and boarding passengers are likely to be inaccurately estimated in the stochastic environment.

The process of passengers arriving, queuing and boarding behavior for vehicles dwelling at stops is analyzed using the PAB process. The process is divided into two consecutive phases. Phase 1 concerns the continuous boarding process and Phase 2 concerns occasions when vehicles wait for passengers, seen hurrying to catch the vehicle before it leaves the stop. Both phases are under the practical constraint of vehicle capacity. Within the small time interval $\tau_i$ (say 1 minute), consider the time for boarding as a renewal process, the mean and variance of passenger boarding time $B(\tau_i)$ can then be derived:

$$\mathrm{E}[B(\tau_i)] = \int_0^\infty x \mathrm{d}G^{N(\tau_i)}(x), \tag{5}$$

and

$$\mathrm{var}[B(\tau_i)] = \int_0^\infty x^2 \mathrm{d}G^{N(\tau_i)}(x^2) + \left[\int_0^\infty x \mathrm{d}G^{N(\tau_i)}(x)\right]^2, \tag{6}$$

where $G^{N(\tau)}(x)$ is the $N(\tau)$th convolution of $G(t)$. $G(t)$ is the cumulative density function of the boarding time for each passenger $B_{per}$, which follows the Normal distribution $B_{per} \sim N(\mu, \sigma^2)$. Note the consistency between the mean of total boarding time and the time interval $\tau_i$ when the number of passenger boarding equals the service rate:

$$\mathrm{E}[B(\tau_i)] = \int_0^\infty x \mathrm{d}G^{N(\tau_i)}(x) = \mathrm{E}[B_{per}^{N(\tau_i)}] = \mathrm{E}[N(\tau_i)] \cdot \mathrm{E}[B_{per}] = \tau_i. \tag{7}$$

Hence, the mean number of passengers boarding at Phase 1 is:

$$n_1 = \sum_i \mathrm{E}[N(\tau_i)] = \sum_i \frac{i}{\mathrm{E}[B_{per}]}. \tag{8}$$

The PAB process is complete in Phase 2. The number of passengers boarding is:

$$n_2 = \lambda(\tau_j) + q'_j, \tag{9}$$

where $q'_j$ is the number of passengers waiting at the beginning of interval $\tau_j$. Applying the same logic as that applied above, the mean and variance of passenger boarding time at this interval is:

$$\mathrm{E}[B(\tau_j)] = \int_0^\infty x \mathrm{d}G^{n_2}(x), \tag{10}$$

$$\mathrm{var}[B(\tau_j)] = \int_0^\infty x^2 \mathrm{d}G^{n_2}(x^2) + \left[\int_0^\infty x \mathrm{d}G^{n_2}(x)\right]^2. \tag{11}$$

The $(n_1 + n_2)$th convolution of $B_{per}$ still follows a Normal distribution, so the total passenger boarding time at stop $s$ of line $l$ follows a Normal distribution: $B_s^l(\tau) \sim N[(n_1 + n_2)\mu, (n_1 + n_2)\sigma^2]$.

During the morning peak hour, when outward trips contribute the heaviest travel demand, the above PAB model can capture the main service process. During afternoon peak, when alighting time occupies the main service time, the alighting time model developed by Adamski (1992) and Lam et al. (1998) can be applied to this model.

## 2.4 Modeling passengers' risk-taking behavior and travel choice

Consider the following general class of passengers: (1) They have a desired arrival time, knowing the travel time is not certain; (2) They will make the travel decisions with $\alpha$ (percent) confidence regarding timely arrival; and (3) They choose the best departure time and transit path as long as the $\alpha$ (percent) confidence of on-time arrival is met. To represent this travel choice and considering reliability requirements, the Chance-constrained model is applied to convert the following stochastic programming problem into a deterministic presentation:

$$\begin{aligned} &\text{Min} \quad t \\ &\text{s.t.} \quad P\{T \le t\} \ge \alpha, \end{aligned} \tag{12}$$

where $T$ is the stochastic travel time.

Classify passengers into $i$ classes. Such passengers are taken to have different confidence levels and are able to introduce different safety margins by a step function:

$$a = sf(\alpha), \quad (\ldots, \alpha_i, \ldots) \in \alpha, \ \alpha \in [0, 1]. \tag{13}$$

*ett* is the effective travel time (ETT) which consists of the mean of passenger travel time and the safety margin:

$$ett_i(\cdot) = \mathrm{E}(C(t,l)) + \phi^{-1}(\alpha_i) \cdot \mathrm{Std}(C(t,l)), \tag{14}$$

where $\alpha_i$ represents the confidence level that $i$th class passengers hold for their on-time arrival requirement. $C(t,r)$ represents passengers' generalized travel cost on route $r$ and departure time $t$. Waiting time consists of (i) passenger waiting time at stops, (ii) the in-vehicle travel cost in unit of time, including passenger waiting time for vehicle departure after boarding, (iii) passenger transfer time (if transfer is needed). Each element is multiplied by a weight coefficient to convert each component to the equivalent unit of time:

$$C(t,r) = \beta_1 T_w(t) + \beta_2 T_v(t) + \beta_3 T_r(t). \tag{15}$$

Passengers generalized travel cost of class $i$ is the summation of ETT, the early and late arrival penalty at destination, and the fares on the transit route. Mathematically, the generalized travel cost is defined as follows:

$$g_i = ett_i + \beta_4 tp(t) + \beta_5 c_f + \varepsilon. \tag{16}$$

The parameters $\beta_4$ and $\beta_5$ are the weight coefficients of early or late penalty and transit fare.

Passenger waiting time is derived from the difference between passenger arrival time $t$ and vehicle arrival time. Denote $\mathbf{T^a_{i,j,l}}$ as the vector of vehicle arrival time at stop $s$ for all vehicles of line $l$ and $\tilde{T}_w$ as delayed waiting time if there is overload delay from the previous vehicle, and $\tilde{T}_w$ equals the headways between two sequential vehicles of the same line.

$$Av_s^l = \min[\mathbf{T^a_{i,j,l}} - t - \tilde{T}_w]^+ \tag{17}$$

represents the earliest arriving vehicle with respect to passenger arrival time $t$ of line $l$ at stop $s$. The passenger waiting time with respect to the vehicle arrival time $t$ is:

$$T_w(t) = Av_s^l - t. \tag{18}$$

This also follows the Normal distribution, drawn from the distribution of vehicle arrival times. The mean and variance of the passenger waiting time can be defined as:

$$\mathrm{E}[T_w(t)] = \mathrm{E}(Av_s^l) - t, \tag{19}$$

$$\mathrm{var}[T_w(t)] = \mathrm{var}(Av_s^l). \tag{20}$$

The in-vehicle travel time and the in-vehicle waiting time can be derived from the vehicle running time model, for the period between vehicle arrival and departure time at stops:

$$T_{i+k,j,l}^a = T_{i,j,l}^d + Tv(t), \tag{21}$$

$$T_{i,j,l}^d = T_{i,j,l}^a + Av_s^l. \tag{22}$$

Passenger transfer cost consists of the waiting time at the transfer stop multiplied by the transfer penalty coefficient $\beta_3$. The early or late penalty is deduced from the passenger departure time, perceived travel time and desired destination arrival time:

$$tp(t) = \begin{cases} \beta'(t^s - \Delta_1^s - ett - t) & \text{if } t^s - \Delta_1^s \ge ett + t \\ \beta''(t + ett - t^s - \Delta_2^s) & \text{if } t^s + \Delta_2^s < ett + t \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

where $[t^s - \Delta_1^s, t^s + \Delta_2^s]$ is the desired arrival time window for the passengers arriving at destination $s$ without any schedule delay penalty. $\beta'$ ($\beta''$) is the unit cost of arriving early or late (i.e. schedule delay) at destinations.

$\varepsilon$ represents passengers perception error when making travel decisions. It is a stochastic variable following the Normal distribution. Probabilities of passengers of class $i$ choosing route $r$ for travel at time $t$ can be expressed as follows:

$$P_i^r(t,l) = \Pr\{g_i(t,l) \le g_i(t',l'), \forall t \ne t', l \ne l'\}, \tag{24}$$

where $l$ is the feasible passenger flow constrained by the available vehicle capacity $l_{\max} : l \in \Omega = \{l | 0 \le l \le l_{\max}\}$.

The stochastic equilibrium condition has been characterized by the following equation (Sheffi 1985):

$$l_r = q \cdot P_r, \tag{25}$$

where $q$ is the average passenger demand for an single OD pair, $l_r$ and $P_r$ are the passenger load and passenger load probabilities of a route connecting the OD pair respectively. Extending the static equilibrium condition for the dynamic and multi-user class network in this paper and representing the variables as vectors, the above stochastic equilibrium condition can be equivalently rewritten as a fixed-point problem:

$$\mathbf{l} - \mathbf{q} \cdot \mathrm{P}(\mathbf{l}) = \mathbf{0}, \tag{26}$$

where $\mathbf{l}$ is the vector of $l_{r,t}^i$, representing passenger class $i$ choosing route $r$ with the departure time $t$; P($\mathbf{l}$) is the vector of $P^i(l_{r,t})$, representing the probability of passenger class $i$ choosing route $r$ with the departure time $t$; and $\mathbf{q}$ is the vector of expected passenger O-D demand.

**Theorem** *At least one solution of the fixed-point problem exists.*

*Proof* The $\Omega$ is a convex and compact set, and $\mathbf{P(l)}$ is continuous on $\Omega$, then follows the Fixed Point Theorem (Gasinski and Papageorgiou 2005), at least one solution exists for the above fixed point problem.

In addition, the regular network flow conservation holds:

$$q^i = \sum_{r,t} q \cdot P^i(l_{r,t}), \tag{27}$$

$$q = \sum_{i} \sum_{r,t} q \cdot P^i(l_{r,t}^i). \tag{28}$$

□

## 3 Dynamic network simulation sub-model and algorithm

The passenger's PAB process is triggered at each time interval by the transit vehicle's arrival at a stop. Passengers are loaded according to the two boarding phases described in the previous section. Passengers loaded at former stops affect the configuration of transit vehicles arriving thereafter. Examples are such that the capacity available and the schedule deviation for downstream stop passengers are determined by service configurations at upstream stops. Thus the service-load dependency is explicitly taken into account throughout the simulation process.

Lam et al. (2008) proposed an algorithm to solve the multi-class reliability-based stochastic user equilibrium (RSUE) problem on road network. This algorithm is adapted to solve the fixed-point model proposed in this paper. Uncertainty effects are recorded in the time-incremental micro-simulation procedure by lines and stops to enable the consideration of boarding delays and on-road schedule deviation. The framework of the solution algorithm is shown in Fig. 3.

## 4 Numerical examples

A simple transit network is used to present the impact of network uncertainties on different passengers' travel behavior. The transit network connects the Kowloon urban area to the Hong Kong International Airport (HKIA) as shown in Fig. 4(a). Four transit lines were considered, the Airport Express Line (AEL), Mass Transit Railway (MTR), Bus line 1 (Bus-1) and Bus line 2 (Bus-2). There are two OD demand pairs connecting the two origins (Kowloon and Tsing Yi) to one destination (HKIA).

Besides passengers being employees working at the airport, passengers go to the airport for multiple purposes, mainly for taking an airplane, picking up passengers or

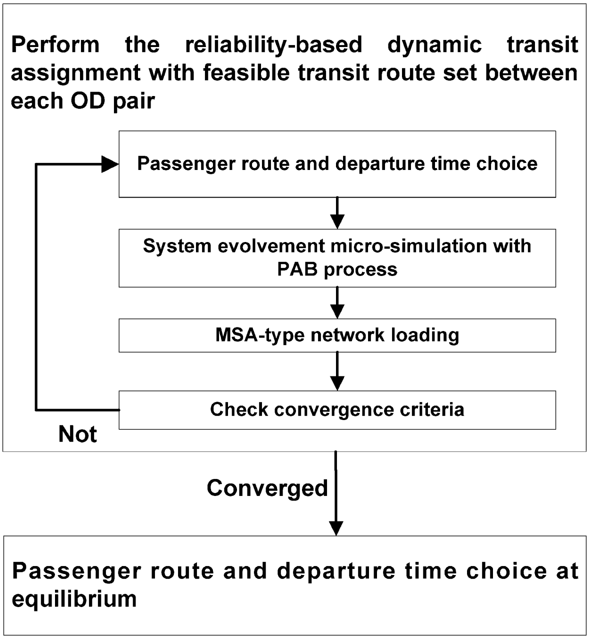


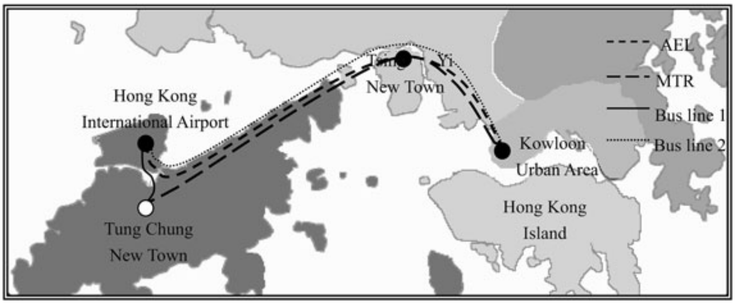

**Fig. 3** The flow chart of the solution algorithm

visiting the museum nearby. Their awareness of trip time and the requirements on trip time reliability are distinguished. The numerical example is designed to: (1) analyze the effects of demand variation on departure time and route choice in the multi-class user network; (2) show how the transit service reliability, by different modes, affects the passenger departure time, route choices and waiting time (3) compare, in terms of average actual travel time and effective travel time, the assignment result under several modeling scenarios.

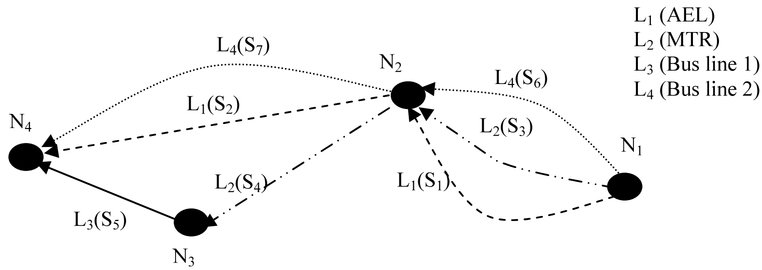
The Hong Kong air flight departure peak period is from 11:00 am to 1:00 pm, and the transit network rush hour to HKIA is around 2 hours prior to flight departures. The transit network study period of the example transit network was thus chosen to be the morning rush period, from 8:00 am to 12:00 noon. Passenger check-in at the airport is usually, 1 hour before flight departure time. Passengers' desired arrival time for this study is therefore set at 11:00 am.

Total passenger demands during the above rush hour have one destination but two origins. They are: (1) from Kowloon (node $N_1$) to HKIA (node $N_4$) $q_{1-4} = 20000$ (pass), and (2) from Tsing Yi (node $N_2$) to HKIA (node $N_4$) $q_{2-4} = 10000$ (pass). Figure 4(b) shows the alternative representation of the example transit network in terms of transit lines and links.

Table 1 gives the basic transit line data for the example transit network. All available transit routes and the attributes of lines are listed in Table 2. AEL is operated strictly according to the given timetable. Owing to exclusive right-of-way operation, AEL and MTR are more reliable than bus lines as the variance of their running times is small. The data provided in Table 1 are from either real data from transit agencies or that gained from practical experience and information systems such as EasyGo




(a) The simplified transit network between Kowloon and HKIA




(b) The alternative representation of the transit network by transit links

**Fig. 4** The transit network of the numerical example

(a research and development product from the Land Surveying and Geo-Informatics Department of Hong Kong Polytechnic University).

The OD demand multiplier is denoted as $\theta$ to represent various passenger demand levels. Other input data include:

the confidence level of passengers $\alpha_1 = 0.5$, $\alpha_2 = 0.7$, $\alpha_3 = 0.95$
the waiting time parameter $\beta_1 = 2$
the in-vehicle travel time parameter $\beta_2 = 1$
the transfer parameter $\beta_3 = 2$
the early or late penalty parameter $\beta_4' = 0.2$, $\beta_4'' = 2$
the fare parameter $\beta_5 = 0.5$
the dispatching headways of AEL, MTR, (BL1) Bus Line 1, and (BL2) Bus Line 2 respectively are: $h_{AEL} = 10$ min, $h_{MTR} = 5$ min, $h_{bus-1} = 4$ min, $h_{bus-2} = 12$ min.

The passenger departure time choices are illustrated in Fig. 5. Different classes of users are included. It is shown that passengers with high reliability requirements will not choose Route 4 (taking the MTR first and then transferring to the bus). This is because of the substantial waiting time uncertainties of transferring and traveling by bus. Passengers demand for Route 1 varies significantly in accordance with different passenger classes. The more expensive yet more reliable route is chosen by risk-averse passengers and only a few risk-neutral passengers make this choice. The indication is that the risk-averse passengers choose the more expensive routes for the sake of reliability. They try to avoid risks by including safety margin in travel de-

**Table 1** Basic transit line data for the example transit network

| Transit line | AEL ($L_1$) | | MTR ($L_2$) | | Bus-1 ($L_3$) | Bus-2 ($L_4$) | |
|---|---|---|---|---|---|---|---|
| $K_l$ (pass/veh) | 500 | | 1500 | | 120 | 120 | |
| Transit link | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_5$ | $S_6$ | $S_7$ |
| In-vehicle time (minutes) Mean / Standard Deviation | 8/ 1 | 12/ 1 | 9/ 1.732 | 11/ 1.732 | 24/ 2.828 | 28/ 3.162 | 35/ 4 |
| Dwell time (Seconds) Mean / Standard Deviation | 23.377/2.241 | | | | –/– (shuttle service) | $2.410 \cdot \textit{Boardings}$/ $1.828 \cdot \textit{Boardings}$ | |
| Transit fare (HK$) | – | 60 | 9 | 14 | 3.5 | 26 | 17 |
| | 90 | | 17 | | | 33 | |

**Table 2** Transit route list by transit links

| Route | Order of transit links | OD pair |
|---|---|---|
| $R_1$ | $S_1$–$S_2$ | $N_1$–$N_4$ |
| $R_2$ | $S_3$–$S_2$ | |
| $R_3$ | $S_3$–$S_7$ | |
| $R_4$ | $S_3$–$S_4$–$S_5$ | |
| $R_5$ | $S_6$–$S_2$ | |
| $R_6$ | $S_6$–$S_7$ | |
| $R_7$ | $S_6$–$S_4$–$S_5$ | |
| $R_8$ | $S_2$ | $N_2$–$N_4$ |
| $R_9$ | $S_7$ | |
| $R_{10}$ | $S_4$–$S_5$ | |

cisions. The extra financial cost to ensure travel time reliability during peak period could be considered an unfair penalty imposed on risk-averse passengers.

It is also observed from Fig. 5 that the risk-neutral passengers and the moderate risk-averse passengers may also choose to travel by Route 4. The number of departing passengers from both these passenger classes increased due to the congestion caused by increased demand. The departure time range did not change for the risk-neutral passengers, but has widened for the moderate risk-averse passengers. The range widened to include extra time at either side of the original departure time. The risk-averse passengers possibly also considered the need to avoid demand driven uncertainties such as vehicle dwell times at stations.

Table 3 shows the proportion of passengers who have been forced to wait at the Kowloon station (i.e. $N_1$) owing to the insufficient capacity in the first arriving. The assignment results of single-class SUE and multi-class RSUE models are compared. The most congested period is the same for both models (from 10:00 to 11:00) from the table. However, the multi-class RSUE model shows an alleviation of overload

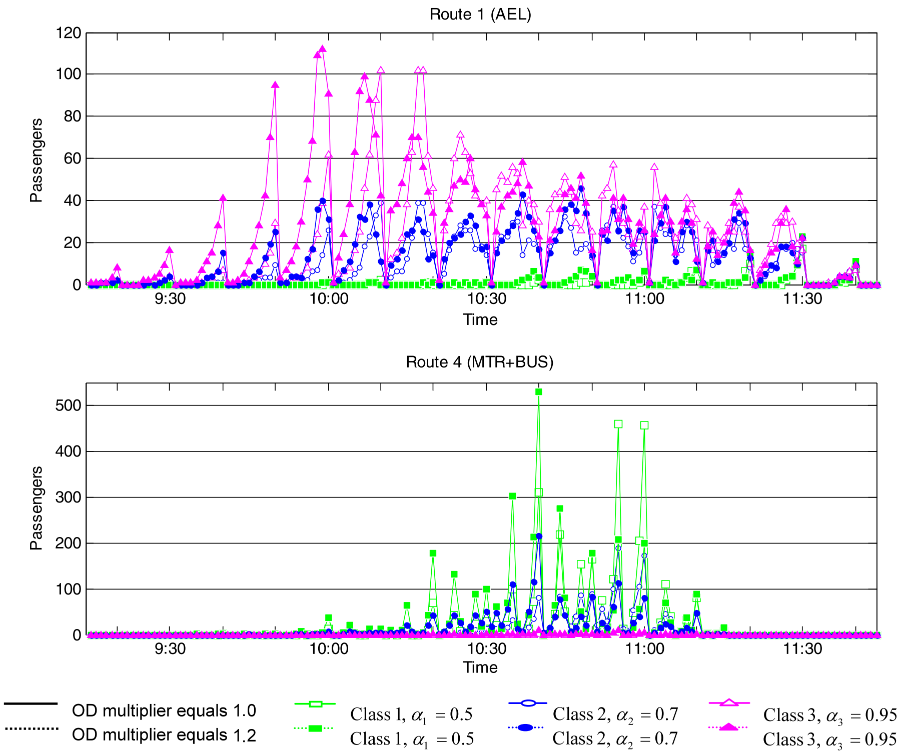



**Fig. 5** Departure time choices of multi-class passengers

**Table 3** Proportion of passengers waiting at $N_1$ resulting from single-class SUE and multi-class RSUE assignment models

| Arrival time | 8:00–8:30–9:00–9:30–10:00–10:30–11:00–11:30–12:00 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Single-class SUE | 0 | 0 | 2% | 9% | **33%** | **16%** | 10% | 2% |
| Multi-class RSUE | 0 | 1% | 8% | 9% | **16%** | **24%** | 11% | 1% |

congestion when the percentage of passengers on the same journey decreased. This decrease may be the result of travel decisions by risk-averse passengers, who changed to earlier departure times or more reliable transit modes.

Table 4 shows passenger total travel cost, effective travel cost, expected travel cost, as well as each cost component of the total travel cost, when the demand multiplier equals 1 and 1.5, respectively. It can be noted that the effective travel time and expected travel time are the same for risk-neutral passengers, which means that network uncertainties have no impact on passenger average in-vehicle travel time and average waiting time. This demonstrates that the SUE model presents a special case of the multi-class RSUE model when network uncertainties are not included in the consideration.

**Table 4** Passengers' travel cost by different classes

| Passenger Classes ($\boldsymbol{\alpha}$) | | OD demand multiplier $\partial = 1$ | | | OD demand multiplier $\partial = 1.5$ | | |
|---|---|---|---|---|---|---|---|
| Parameters | | 0.5 | 0.7 | 0.95 | 0.5 | 0.7 | 0.95 |
| Total travel cost (min) | | 119.36 | 118.50 | 107.79 | 338.95 | 256.55 | 162.49 |
| Effective travel time (min) | | 61.51 | 60.54 | 47.042 | 122.89 | 115.36 | 73.53 |
| Expected travel time (min) | | 61.51 | 54.85 | 43.06 | 122.89 | 109.88 | 69.21 |
| Money cost (%) | | 9% | 20% | **35%** | 3% | 11% | **27%** |
| Early or late penalty (%) | | **41%** | **37%** | 27% | **57%** | **43%** | **27%** |
| In-vehicle time (%) | | **29%** | 15% | 10% | 13% | 11% | 12% |
| Waiting time (%) | | 21% | **28%** | **28%** | **27%** | **35%** | **34%** |
| Standard | In-vehicle | 4.09 | 2.58 | 1.52 | 4.08 | 2.51 | 1.55 |
| deviation (min) | Waiting | 1.56 | 0.73 | 0.69 | 1.58 | 0.82 | 0.77 |

However, the risk-neutral passengers have the highest total travel cost for both the above scenarios (demand multiplier equals 1 and 1.5). The highest component of their total travel cost is the early or late penalty. This implies that these passengers do not have the sufficient recognition of network uncertainties and the possibility of unreliable travel time, consequently no safety margin is added to their expected travel time. As a result, such passengers may depart during the most congested time period (35% waiting time when $\partial = 1.5$) or choose the time-consuming route (29% in-vehicle travel time when $\partial = 1$) for travel.

Of the three passenger classes, the strongly risk-averse passengers are subjected to the least total travel cost but the highest financial cost, compared with other passenger classes. The highest financial cost for this class is likely to be the result of the willingness of risk-averse passengers to pay extra money to ensure travel time reliability. Thus, their cost of early or late penalty is the lowest among three passengers classes (takes 27% of total cost), which means that the reliability of on-time arrival was maintained. The highest cost for moderate risk-averse passengers lies in early or late penalties and waiting time, indicating the possibility of overload delays as the result of later departures.

Vehicle load with deterministic capacity under demand uncertainties are illustrated in Fig. 6. It is seen that the mean number of passengers on-board varies considerably with the different overload parameters shown. When the overload parameter equals 0.1 (which means transit vehicles are not usually congested), the full stochastic capacity, plus a possible one standard deviation passenger variation, could be loaded without causing any in-vehicle congestion. However in extremely congested situations when the overload parameter equals 0.9 (which means vehicles are usually congested out of 90% situations), setting the design vehicle capacity to bind the practical loaded passengers will lead to the underestimate the number of passengers on-board. When the overload parameter equals 0.5, vehicles are considered full for 50% situations, but still have another half probabilities of being overcrowded. This situation indicates that use only the design capacity in the stochastic network will cause severe underestimation of passenger vehicle load and as such lead to a degraded level of service.

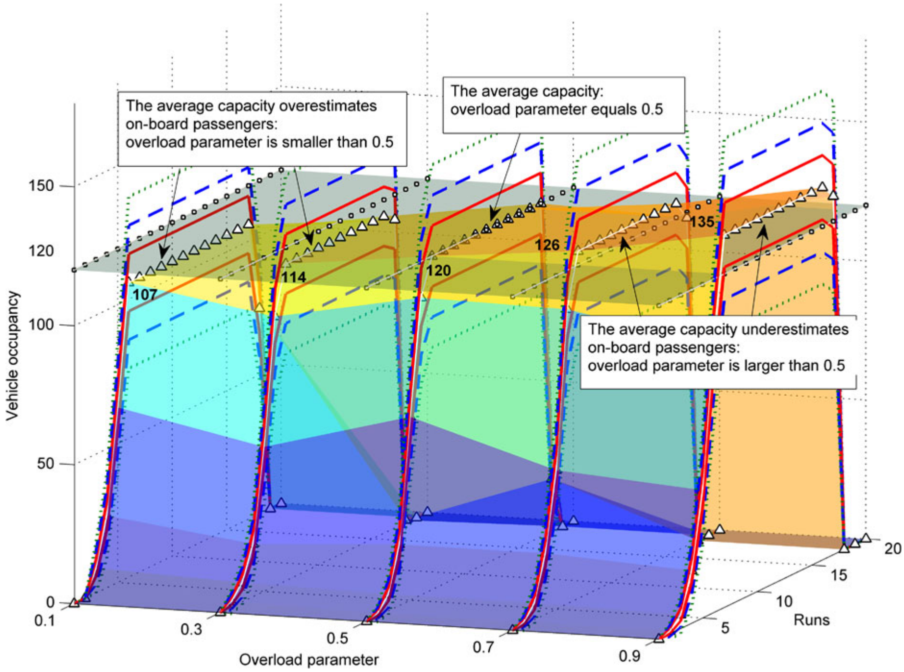



**Fig. 6** Passenger load with vehicle capacity constraint under passenger demand uncertainties

## 5 Conclusions

A new reliability-based dynamic transit assignment model has been presented in this paper to investigate multi-class passengers travel decisions including route and departure time choices in congested and stochastic transit networks with uncertainties. The proposed model has been shown to be capable of accounting for the impact of uncertainties with respect to passenger travel time, waiting time, vehicle dwelling time, as well as transit service reliability. It is noted that the SUE model presents only a special case (passengers are risk-neutral) of the multi-class RSUE model. The results of the assignment showed on the same transit network, risk-neutral passengers have the highest travel cost, in comparison with that of risk-averse passengers. Risk-averse passengers are likely to choose the less risky options, regardless of higher money cost. They disregard the total travel cost saving from early or late penalty and in-vehicle travel time and are willing to spend more money in the hope of benefits of reliability.

This study also demonstrated the importance of knowledge of demand and demand-driven uncertainties in transport networks and the effect they may have on passenger departure time choices and the corresponding reliability of reaching a chosen destination on time. Without such knowledge transport planners would have difficulty providing a transit service of quality and in the ever changing situations of any city, such ignorance could lead to a rapid degrading of transit services.

This research study has also provided a new fundamental tool for transit service evaluation and transit network design under demand and supply uncertainties in the schedule-based network framework. Further research issues should include:

(i) Calibration of the stochastic parameters, such as boarding time per passenger, passenger arrival and OD demand distributions.
(ii) Design of transit services for schedule coordination under transit network uncertainties due to adverse weather and traffic incidents.

**Acknowledgements** The work described in this paper was jointly supported by a grant from the Research Grant Council of the Hong Kong Special Administrative Region to the Hong Kong Polytechnic University (Project No. 5202/06E). The first author would like to thank The Hong Kong Polytechnic University for support in carrying out this MPhil research in Hong Kong under the supervision of the second and third authors.

## References

Adamski A (1992) Probabilistic models of passengers service processes at bus stops. Transp Res 26B:253–259
Bell MGH, Schmöcker JD, Iida Y, Lam WHK (2002) Transit network reliability: an application of absorbing Markov chains. In: Taylor MAP (ed) Proceedings of the 15th international symposium on transportation and traffic theory. Elsevier Science, Amsterdam, pp 43–62
Chen A, Ji Z, Recker W (2002) Travel time reliability with risk-sensitive travelers. J Transp Res Board 1873:27–33
Cominetti R, Correa J (2001) Common lines and passenger assignment in congested transit networks. Transp Sci 35:250–267
De Cea J, Fernandez E (1993) Transit assignment for congested public transport system: an equilibrium model. Transp Sci 27:133–147
Gasinski L, Papageorgiou NS (2005) Nonlinear analysis. Chapman & Hall/CRC, London
Hamdouch Y, Lawphongpanich S (2008) Schedule-based transit assignment model with travel strategies and capacity constraints. Transp Res 42B:663–684
Hickman MD (2001) An analytic stochastic model for the transit vehicle holding problem. Transp Sci 35:215–237
Kurauchi F, Bell MGH, Schmöcker JD (2003) Capacity constrained transit assignment with common lines. J Math Model Algorithms 2:309–327
Lam WHK, Cheung CY, Poon YF (1998) A study of train dwelling time at the Hong Kong mass transit railway system. J Adv Transp 32:285–296
Lam WHK, Shao H, Sumalee A (2008) Modeling impacts of adverse weather conditions on a road network with uncertainties in demand and supply. Transp Res 42B:890–910
Larrain H, Muñoz JC (2008) Corridor assignment assuming congestion due to passenger boarding and alighting. Netw Spat Econ 8:241–256
Lo H, Luo XW, Siu BWY (2006) Degradable transport network: travel time budget of travelers with heterogeneous risk aversion. Transp Res 40B:792–806
Noland RB, Polak JW (2002) Travel time variability: a review of theoretical and empirical issues. Transp Rev 22:39–54
Nuzzolo A, Russo F, Crisalli U (2001) A doubly dynamic schedule-based assignment model for transit networks. Transp Sci 35:268–285
Nuzzolo A, Russo F, Crisalli U (2003) Transit network modeling: the schedule-based dynamic approach. Franco Angeli srl, Milana, Italy
Poon MH, Wong SC, Tong CO (2004) A dynamic schedule-based model for congested transit networks. Transp Res 38B:343–368
Powell WB, Sheffi Y (1983) A probabilistic model of bus route performance. Transp Sci 17:376–404
Schmocher JD, Bell MGH, Kurauchi F (2008) A quasi-dynamic capacity constrained frequency-based transit assignment model. Transp Res 42B:924–945

Sheffi Y (1985) Urban transport networks: equilibrium analysis with mathematical programming methods. Prentice-Hall Inc, Englewood Cliffs

Spiess H, Florian M (1989) Optimal strategies: a new assignment model for transit networks. Transp Res 23B:83–102

Szeto WY, Solayappan M Wong KI (2009) Reliability-based stochastic transit assignment with capacity constraints. CD-ROM of 11th international conference on advanced systems for public transport, Hong Kong SAR

Teklu F, Watling D, Connors R (2007) A Markov process model for capacity constrained transit assignment. In: Transportation and traffic theory, pp 483–506

Wilson N, Nuzzolo A (2004) Schedule-based dynamic transit modeling—theory and application. Kluwer Academic, Dordrecht

Wirasinghe SC (2003) Initial planning for urban transit systems. In: Advanced modeling for transit operations and services planning. Elsevier, Amsterdam

Yang L, Lam WHK (2006) Probit-based reliability-based transit assignment model. Transp Res Rec 1977:154–163

Zhang YQ, Lam WHK, Sumalee A (2009) Dynamic transit assignment model for congested transit networks with uncertainties. CD-ROM of 88th transportation research board meeting, Washington, DC