Dynamic Journey Time Estimation in Stochastic Road Networks with Uncertainty

Qiong TANG a, Xingang LI b, William H.K. LAM c*, H.W. HO d

a,b,c,d Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
a Department of Economics and Management, Hengyang Normal University, Hengyang, 421002, China
b,c School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China.
a Email: tangqiong_1985@163.com 
b Email: lixingang@bjtu.edu.cn 
c Email: william.lam@polyu.edu.hk 
d Email: cehwho@polyu.edu.hk

Abstract: This study investigates the dynamic journey time estimation (DJTE) problem, which can be used to simultaneously estimate the mean and standard deviation (SD) of path journey times in stochastic road networks. A bi-level programming model is proposed to solve the DJTE problem. The upper level is used to minimise the deviations of the mean and SD between the estimated and observed path journey times. The lower level is a reliability-based dynamic traffic assignment model taking into account the variation of path journey times. The modified simulated annealing (SA) algorithm is adapted to solve the proposed bi-level problem. A numerical example is used to illustrate the application of the proposed DJTE model and the modified algorithm together with insightful findings.

Keywords: Dynamic journey time; Dynamic traffic assignment; Simulated annealing algorithm.

1. INTRODUCTION

Journey time (travel time) is an important measure of the performance of a road network, so accurate and reliable travel time information is increasingly important for traffic engineers. Travel time is a common measure for road users, enabling them to make travel choices and to avoid unnecessary delays (Liu and Ma, 2009). Accurate estimations of travel times are important for improving traffic operations and identifying key bottlenecks in a traffic network (Zhan et al., 2013). Various models have been proposed to contribute to the field of journey time estimation, which can be classified into two categories: static models for long-term purposes (Lam et al., 2002; 2006; Shao et al., 2013; Zhan et al., 2013; Jenelius and Koutsopoulos, 2013; Uchida, 2014) and dynamic models for short-term purposes (Dion and Rakha, 2006; Liu et al., 2007; Kamga et al., 2011; Sumalee et al., 2013). Static models focus on the static description of traffic flows in the network, implying that traffic flow and travel times are invariant over the duration of the peak period. The dynamic journey time estimation (DJTE) model developed in this study is in the latter category and is proposed in response to the untenable assumption that traffic flow is static. In contrast to the estimations of average travel

* Corresponding author.
times over the entire peak period in static models, in this model the travel times for travellers departing at different times can be estimated separately.

To estimate travel times accurately within a road network, the models use different types of traffic data collected by various technologies: electronic distance-measuring instruments, electronic license plate matching, cellular phone tracking, automatic vehicle identification (AVI), automatic vehicle location, global positioning systems (GPS), probe vehicles, virtual probe vehicles and wireless magnetic sensors (Shao et al., 2013). In general, AVI is a proven technology for providing area-wide real-time traffic data with low operating costs. AVI technology is suitable for annual, daily and real-time traffic monitoring (Tam and Lam, 2008), and is thus used in this study for collecting real-time link traffic flows and path travel times.

Travel time estimation in traffic networks is challenging as it is intrinsically uncertain (Zhang and Zuylen, 2013). This uncertainty occurs on both the supply side (e.g., weather conditions, traffic incidents and capacity variation) and the demand side (e.g., population characteristics, traffic information and demand fluctuation) of the network. Travellers find it difficult to predict exactly when they will arrive at their destination. Hence, evaluating only the mean of travel times may not be sufficient for capturing their variability and reliability, so in this study a model for simultaneously estimating the mean and variance of path travel times is proposed, to effectively identify variability. The mean and variance measures of travel time can reasonably represent travel time variability, both in theory and in practice (Jackson and Jucker, 1981). Jackson and Jucker (1981) found from their empirical study that perceived travel time reliability is an important component of the traveller’s route choice decision, and suggested that including the variability of travel time in the impedance function with mean travel time might improve the traffic assignment process.

Hall (1983) suggested that travellers tend to reserve a safety margin in their travel time to mitigate the unreliability, improving their likelihood of arriving on time. Lam and Small (2001) also reported in their empirical study that travellers are likely to set up a travel time safety margin to avoid late arrival. Travel time and reliability undoubtedly influence travellers’ route choice. Many models have been proposed (Lo et al., 2006; Shao et al., 2006; Chen et al., 2012) to capture risk-based route choice behaviour, using the concept of effective travel time or travel time budget, defined as the mean travel time plus a safety margin. With this definition, these models ensure that the probability of completing the trip within the effective travel time or travel time budget is not less than the predefined reliability threshold or confidence level $\alpha$. The effective travel time is therefore used in this study as the route choice criterion of travellers. In the proposed DJTE model, a dynamic traffic assignment is used to model travellers’ route choice behaviour in dynamic networks.

Lam et al. (2002, 2006) developed a traffic flow simulator based on a probit-type stochastic user equilibrium (SUE), to estimate travel times by minimising the deviation between the observed and estimated link flows and origin-destination (OD) demands. Shao et al. (2013) proposed a journey time estimator to estimate the stochastic journey time by taking into account the reliability-based stochastic user equilibrium traffic assignment problem. Zhan et al. (2013) presented a new descriptive model for estimating the hourly average of urban link travel times using taxicab OD trip data, by minimising the error between the estimated and observed path travel times. Jenelius and Koutsopoulos (2013) developed a statistical model for urban road network travel time estimation based on low-frequency GPS probe vehicle data. Their estimates consist of running times on links and delays at traffic signals/intersections. Uchida (2014) proposed a utility maximisation framework for the simultaneous estimation of the value of travel time and its reliability based on a risk-averse driver’s route choice behaviour. These studies (Lam et al. 2002; 2006; Shao et al., 2013; Zhan et al., 2013; Jenelius and Koutsopoulos, 2013; Uchida, 2014) all fall within the static model category. Their disadvantages are that they
do not capture the time-dependent characteristics of traffic flow and they assume the observed link flows represent a steady-state situation that persists over a block of time (Sherali and Park, 2001).

Most dynamic travel time estimation models use the mean travel times as the route choice criterion but ignore their variances. Kamga et al. (2011) presented a methodology to estimate travel time using a simulation-based dynamic traffic assignment model. Sumalee et al. (2013) proposed a framework for evaluating the distributions of stochastic dynamic link travel time and journey time, which also assesses journey time reliability. However, both the mean and variance of travel time should be included as part of the impedance function in both the route choice and the modelling process (Jackson and Jucker, 1981). One advantage of the model proposed in this study over the abovementioned dynamic models is that the mean and variance of travel time are estimated by taking into account the traveller’s route choice behaviour in a road network. Liu et al. (2007) estimated dynamic travel times by minimising the summation errors of the mean and SD between the observed and estimated travel times, and used a mixed logit formulation to calculate travellers’ route choice probability. Using this method to capture travellers’ route choice is certainly a calculation convenience, but it is not suitable for a dynamic network.

The covariance relationship is also crucial in the travel time estimation process. The covariance relationships of link travel times can represent the magnitude of the dependency of travel times on two alternative links. This dependency can provide useful information for path travel time estimation (Shao et al., 2013). Lam et al. (2002) concluded that the covariance information has a significant influence on the accuracy of travel time forecasting. Xu (2003) conducted sensitivity tests and found that covariance information can improve the travel time estimation, particularly for congested road networks.

Similar to the travel time estimation problem presented by Shao et al. (2013), the DJTE proposed in this paper can be formulated as a bi-level problem. The upper-level problem is a generalised least-squares optimisation problem that minimises the deviations of the mean and SD between the observed and estimated path travel times. The lower-level problem formulates travellers’ route choice behaviour as a reliability-based dynamic traffic assignment (DTA) problem, which considers the covariance information of link travel times. This study includes the following three features: (1) The mean and SD of path travel times are simultaneously estimated by considering travellers’ route choice; (2) travel time and its reliability are incorporated into the DTA and (3) covariance information of link travel times is taken into account in the travel time estimation process.

The remainder of this paper is organised as follows. In the next section, the notation and bi-level model formulation of the DJTE problem is described. In Section 3, a simulated annealing-based solution algorithm is proposed to solve the bi-level problem. Numerical examples of a simple small network are depicted in Section 4, in which the application of the DJTE and the efficiency of the proposed solution algorithms are demonstrated. The conclusions and further studies are given in Section 5.

2. MODEL FORMULATION

2.1 Notations and Assumptions

The notations used throughout this paper are listed in Table 1 unless otherwise specified.
Table 1. List of variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G = (N, A)$</td>
<td>A network composed of a set of nodes $N$, and a set of links $A$</td>
</tr>
<tr>
<td>$r$</td>
<td>Origin node</td>
</tr>
<tr>
<td>$s$</td>
<td>Destination node</td>
</tr>
<tr>
<td>$R$</td>
<td>Set of OD pairs</td>
</tr>
<tr>
<td>$P_{rs}$</td>
<td>Set of paths between OD pair $(r, s)$</td>
</tr>
<tr>
<td>$\bar{A}$</td>
<td>Observed link set in which the traffic flow can be observed</td>
</tr>
<tr>
<td>$\bar{P}_{rs}$</td>
<td>Observed path set in which the travel time can be observed</td>
</tr>
<tr>
<td>$T_d$</td>
<td>Set of the departure time intervals</td>
</tr>
<tr>
<td>$T_o$</td>
<td>Set of the observed time intervals $T_o \subseteq T_d$</td>
</tr>
<tr>
<td>$C_a$</td>
<td>Capacity of link $a \in A$</td>
</tr>
<tr>
<td>$T_a$</td>
<td>Free flow travel time of link $a \in A$</td>
</tr>
<tr>
<td>$k$</td>
<td>Subscript for the departure time intervals $k \in T_d$</td>
</tr>
<tr>
<td>$t$</td>
<td>Subscript for the observation time intervals $t \in T_o$</td>
</tr>
<tr>
<td>$x_a(k)$</td>
<td>Link flow on link $a \in A$ at the beginning of time interval $k$</td>
</tr>
<tr>
<td>$u_a(k)$</td>
<td>Link inflow rate to link $a \in A$ at the beginning of time interval $k$</td>
</tr>
<tr>
<td>$v_a(k)$</td>
<td>Link exit flow rate from link $a \in A$ at the beginning of time interval $k$</td>
</tr>
<tr>
<td>$T_a(k)$</td>
<td>Travel time of link $a \in A$ at the end of time interval $k$</td>
</tr>
<tr>
<td>$\sigma(k)$</td>
<td>Mean of travel time of link $a \in A$ at the end of time interval $k$</td>
</tr>
<tr>
<td>$\sigma_a(k)$</td>
<td>SD of travel time of link $a \in A$ at the end of time interval $k$</td>
</tr>
<tr>
<td>$f_p^rs(k)$</td>
<td>Flow on path $p \in P_{rs}$ at the beginning of time interval $k$</td>
</tr>
<tr>
<td>$T_p^rs(k)$</td>
<td>Travel time on path $p \in P_{rs}$ at the end of time interval $k$</td>
</tr>
<tr>
<td>$\mu_p^rs(k)$</td>
<td>Mean of travel time on path $p \in P_{rs}$ at the end of time interval $k$</td>
</tr>
<tr>
<td>$\sigma_p^rs(k)$</td>
<td>SD of travel time on path $p \in P_{rs}$ at the end of time interval $k$</td>
</tr>
<tr>
<td>$\pi_p^rs(k)$</td>
<td>Effective path journey time on path $p \in P_{rs}$ at the end of time interval $k$</td>
</tr>
<tr>
<td>$\delta_{aq}^rs(l)$</td>
<td>Minimum effective path journey time on path $p \in P_{rs}$ at the end of time interval $k$</td>
</tr>
<tr>
<td>$\delta_{aq}^rs(l)$</td>
<td>Minimum effective path journey time on path $p \in P_{rs}$ at the end of time interval $k$</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Tolerance for the estimated link flows deviated from the observed link flows</td>
</tr>
<tr>
<td>$\lambda_{rs}$</td>
<td>Demand multiplier to the prior mean OD demand $d^rs$ for $rs \in R$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Vector of $(\cdots \lambda_{rs} \cdots)^T$ for $rs \in R$</td>
</tr>
</tbody>
</table>
To facilitate the presentation of the essential ideas without loss of generality, the following basic assumptions are made.

A1. Journey time variations result from the road supply uncertainty. The observed data under demand uncertainty are excluded from the analysis. The time-dependent travel demands between each OD pair are assumed to be given and deterministic, as the travel demand considered here is short-term and within the day (Corthout et al., 2010; Knoop et al., 2010; Kamga et al., 2011).

A2. The stochastic dynamic link performance function is used for the flow propagation component to capture the stochastic effects. Empirical results show that traffic parameters; free flow speed, capacity, etc., can vary greatly due to the variability in driving behaviour and the characteristics of vehicles (Wang et al., 2013). Similar to Lam et al. (2008), the free flow travel time and capacity are assumed to be random parameters in the link performance function.

A3. Travellers make path choice decisions based on the effective path travel time. The user equilibrium based on effective travel time or travel time budget is an extension of the classic Wardrop equilibrium, and has been used in static models (Lo et al., 2006; Shao et al., 2006; Chen et al., 2012). In this study, it is further extended to dynamic networks. The covariance between link travel times is considered in calculating the effective path travel time.

A4. The variance of the link travel time is an increasing function with respect to the mean of the link travel time.

### 2.2 Definition of Effective Path Travel Time

We choose the following linear form for the link travel time function:

\[
T_a(k) = 
T_a^{f} + \omega_a \frac{x_a(k)}{C_a} \tag{1}
\]

where \( \omega_a \) is the congestion-dependent coefficient in the link travel time function. \( T_a^{f} \) and \( C_a \) are free flow travel time and capacity, respectively. These are random variables that describe the stochastic characteristics of network supply uncertainties. Similar to previous studies (e.g., Lam et al., 2008), it is assumed that \( T_a^{f} \) and \( 1/C_a \) are independent random variables with normal distributions. The mean of \( T_a^{f} \) and \( 1/C_a \) are respectively taken as \( \mu_a^{f} \) and \( \mu_a^{C} \). According to Equation (1) and assumption A4, the mean and SD of the stochastic link travel time can be expressed as follows:

\[
t_a(k) = \mu_a^{f} + \omega_a x_a(k) \mu_a^{C} \tag{2}
\]

\[
\sigma_a(k) = \beta t_a(k) \tag{3}
\]

where \( \beta \) is a non-negative decision variable to be estimated by DJTE. Assume a path \( p \) between OD pair \((r,s)\) consists of the following set of links \( \{a_1, a_2, \ldots, a_m\} \) and nodes \((r,1,2,\cdots,m-1,s)\). The path travel time \( T_p^\alpha(k) \) follows a multivariate normal distribution. The mean \((t_p^\alpha(k))\) and...
SD \( (\sigma_{rs}^p(k)) \) of the stochastic path travel time \( T_{rs}^p(k) \) for this path \( p \) can be calculated as

\[
  t_{rs}^p(k) = \sum_{a \text{ on path } p \mid (2k) \in T_a} t_a(l)\delta_{apk}^rs(l)
\]

\[
  \sigma_{rs}^p(k) = \sqrt{\left(\sum_{a \text{ on path } p \mid (2k) \in T_a} \sigma_a(l)\delta_{apk}^rs(l)\right)^2 + 2 \sum_{1 \leq i < j \leq m} \text{cov}(t_a, t_d)}
\]

where \( \text{cov}(t_a, t_d) \) is the covariance between travel time for link \( a \) and \( d \) on path \( p \). For example, a positive covariance between travel time for link \( a \) and \( d \) indicates that a longer travel time on link \( a \) corresponds to a long travel time on link \( d \).

As mentioned, it is realistic to assume that travellers include an extra travel time budget due to the variations of path travel time, guaranteeing on-time arrival with a desired probability while making their path choice. Similar to previous research (Lo et al., 2006; Chen et al., 2012) we use the following expression to define the effective path journey time required to traverse path \( p \) with a given probability \( \alpha \).

\[
  \eta_{rs}^{p,\alpha}(k) = t_{rs}^p(k) + Z_\alpha \sigma_{rs}^p(k)
\]

where \( Z_\alpha \) is the inverse of the cumulative distribution function of standard normal distribution at \( \alpha \) confidence level, which can be predetermined from the travellers’ trip purpose. For example, if \( Z_\alpha \) is taken as 1.64 in Equation (6), travellers have a 95\% chance (i.e., \( \alpha = 95\% \)) of arriving at their destination within the effective path journey time calculated in Equation (6).

To calculate the value of \( \eta_{rs}^{p,\alpha}(k) \), the link travel time \( T_a(k) \), which can be calculated by the traffic pattern \( f \) (such as link traffic flows \( x_a(k) \), \( \delta_{apk}^rs(l) \)), and link travel time covariance \( \text{cov}(t_a, t_d) \), should first be obtained. The traffic pattern can be obtained by solving the dynamic traffic assignment problem, which will be discussed in detail in the following section. For link travel time covariance, according to Chen et al. (2012), there are no empirical findings on the temporal correlation between link travel times. Thus, we only consider the spatial covariance among link travel times. For simplicity, the covariance between \( t_a(k_1) \) and \( t_b(k_2) \) is expressed as:

\[
  \text{cov}(t_a(k_1), t_b(k_2)) = \theta \exp(-|\text{order}(a) - \text{order}(b)|)\sigma_a(k_1)\sigma_b(k_2)
\]

where link \( a \) and \( b \) are on the same path \( p \) and \( k_1, k_2 \in T_a \); the function \( \text{order}(a) \) is the ordinal number of link \( a \) on the path \( p \); \( \theta \) has a deterministic value with \( \theta > 0 \) (\( \theta < 0 \)) indicating that the link travel times are positively (negatively) correlated. For \( \theta = 0 \), there is no correlation between link travel times.

### 2.3 DTA Based on an Effective Path Travel Time

If all travellers choose a path based on an effective path journey time, the dynamic user equilibrium condition implies that at each time interval \( k \) any used path has an identical and minimum effective path travel time (Equation 8):
\[ n_p^{rs}(k) \begin{cases} = \pi_p^{rs}(k), & \text{if } f_p^{rs}(k) > 0 \\ \geq \pi_p^{rs}(k), & \text{if } f_p^{rs}(k) = 0 \end{cases} \] (8)

\[ D^r(k) = \sum_{p \in P_r} f_p^{rs}(k), \forall r, s, k \] (9)

\[ u_s(k) = \sum_{p \in P_r} \delta_{ap} f_p^{rs}(k), \forall r, s, a, p \] (10)

where Equation (9) and (10) is the path flow assignment constraints for the DTA. It is assumed that the estimated OD demand is in proportion to the prior OD demand, i.e., \( D^r(k) = \lambda_{rs} d^r(k) \), where \( d^r(k) \) is given a prior mean OD demand and \( \lambda_{rs} \) is the demand multiplier, which is a decision variable to be estimated by the DJTE. Similar to the previous DTA problems (Huang and Lam, 2002; Yin et al., 2004; Long et al., 2011, 2013a, 2013b), the flow conservation and propagation equations (Equations 11 ~ 13), definitional constraints (Equation 14) and non-negativity conditions (Equations 15 and 16) for this DTA problem are defined as follows:

\[ \sum_{a \in A_B(j)} u_{a,p}^{rs}(k) = \sum_{a \in A_B(j)} v_{a,p}^{rs}(k), \forall r, s, p; \forall j \neq r, s \] (11)

\[ \frac{dx_{a,p}^{rs}(k)}{dt} = u_{a,p}^{rs}(k) - v_{a,p}^{rs}(k), \forall r, s, a, p \] (12)

\[ v_{a,p}^{rs}(k + t_{a,p}^{rs}(k)) = \frac{u_{a,p}^{rs}(k)}{1 + dt_{a,p}^{rs}(k) / dk}, \forall r, s, a, p \] (13)

\[ \sum_{a \in A_B(j)} u_{a,p}^{rs}(k) = u_a(k), \sum_{a \in A_B(j)} v_{a,p}^{rs}(k) = v_a(k), \sum_{a \in A_B(j)} x_{a,p}^{rs}(k) = x_a(k), \forall a \] (14)

\[ u_{a,p}^{rs}(k) \geq 0, v_{a,p}^{rs}(k) \geq 0, x_{a,p}^{rs}(k) \geq 0, \forall r, s, a, p \] (15)

\[ f_p^{rs}(k) \geq 0, \forall r, s, p, k \] (16)

According to the above equations, the mean and SD of the path travel time can be written as a function of \( \lambda, \beta, f \) and \( k \) as follows.

\[ t_p^{rs}(k) = t_p^{rs}(\lambda, \beta, f, k) \] (17)

\[ \sigma_p^{rs}(k) = \sigma_p^{rs}(\lambda, \beta, f, k) \] (18)

### 2.4 Upper-Level Model

The upper-level model aims to determine \( \lambda \) and \( \beta \). A generalised least-squares function is used as the objective function so that the deviations between the observed and estimated mean and SD of the path travel time are minimised. Apart from the observed travel times, the observed link flows are treated as the constraints in this upper-level model, formulated as follows.

\[ \min_{\lambda, \beta} z_i = \sum_{r \in R} \sum_{p \in P_r} \sum_{s \in R} (\tilde{t}_p^{rs}(\lambda, \beta, f, t) - \bar{t}_p^{rs}(t))^2 + \sum_{r \in R} \sum_{p \in P_r} (\tilde{\sigma}_p^{rs}(\lambda, \beta, f, t) - \bar{\sigma}_p^{rs}(t))^2 \] (19)

subject to
The objective function of Equation (19) is to minimise the weighted sum of the deviation between the estimated and observed mean and SD of the path travel time. Constraint (20) ensures the estimated link flows are within a reasonable range of observed link flows. Constraints (21) and (22) are boundary constraints of the decision variance $\lambda$ and $\beta$, respectively.

### 2.5 Lower-Level Model

The discrete time version of a DTA model that is based on an effective path journey time can be formulated as an equivalent variational inequality (VI) problem for finding a vector $f^{*} \in F$, so that for all $f \in F$, Equation (23) is satisfied.

$$\sum_{k \in D_{D}, p \in P_{r}} \eta^{r,s*}(k)[f^{rs*}(k) - f^{rs}(k)] \geq 0$$

where $F$ is a closed convex set defined by

$$F = \{ f \geq 0 : \sum_{p \in P_{r}} f^{rs}(k) = D^{rs}(k), \forall rs \in R, k \in T_{d} \}$$

By considering the upper- and lower-level models described in Sections 2.4 and 2.5, the lower-level problem is clearly to determine the traffic pattern $f$ with fixed $\lambda$ and $\beta$ given by the upper level. With traffic pattern $f$ from the lower-level model, the optimal valued $\lambda$ and $\beta$ must be found in the upper-level problem.

### 3. SOLUTION ALGORITHM

In general, the bi-level DJTE problem is solved as follows. For an initial $\lambda$ and $\beta$ in the upper level, traffic pattern $f$ can be obtained through solving DTA in the lower-level model. With the traffic pattern $f$ from the lower-level model and the initial value of $\lambda$ and $\beta$, the objective function of the upper level can be calculated and a new set of $\lambda$ and $\beta$ estimated. The above procedure is then repeated until it converges to the optimal solution of the bi-level model. In this section, the method to obtain the traffic pattern $f$ by solving the DTA will first be presented, and solving the bi-level programming model using the SA approach will be illustrated.

#### 3.1 Solution Algorithm for DTA

In the effective path journey time calculation process, link-based algorithms cannot be used to solve the DTA model due to the non-additive properties. Thus, a path-based algorithm based on the method of successive averages (MSA) and the column generation method for updating the route set is used.
Step 1 **Initialisation**: Set all link flows $x_{rs}^n(k), u_{rs}^n(k), v_{rs}^n(k), \forall k \in T_d$ to zero and calculate the initial link travel time $t_{rs}^0(k), \forall k \in T_d$. Set the iteration counter to $n=1$, the maximum iteration to $N$, and the convergence criterion to $\varepsilon$.

Step 2 **Shortest path**: Find the shortest path $p_{rs}^{\alpha,\beta}(k)$ for each OD pair at each time interval and calculate the corresponding $\pi_{rs}^{\alpha,\beta}(k)$. If the shortest path is new, then update the path set $P$.

Step 3 **Flow assignment**: Assign all OD demand on the shortest path based on effective path travel time, then get $f_{rs}^{\alpha,\beta}(k)$. Calculate the new path flow

$$f_{rs}^{\alpha,\beta}(k) = f_{rs}^{\alpha,\beta-1}(k) + \frac{1}{n+1}[f_{rs}^{\alpha,\beta}(k) - f_{rs}^{\alpha,\beta-1}(k)]$$

Step 4 **Network flow loading**: Update all link flows $x_{rs}^n(k), u_{rs}^n(k), v_{rs}^n(k)$ and the link travel time $t_{rs}^n(k)$, based on the new path flow in Step 3. Then calculate the effective path journey time $\eta_{rs}^{\alpha,\beta}(k)$ considering the link travel time covariance.

Step 5 **Convergence check**: If

$$\sum_{rs} \left| \pi_{rs}^{\alpha,\beta}(k)f_{rs}^{\alpha,\beta} - \pi_{rs}^{\alpha,\beta}(k)f_{rs}^{\alpha,\beta} \right| \leq \varepsilon$$

or $n=N$, stop; otherwise, $n=n+1$, go to Step 2.

### 3.2 Solution Algorithm for Bi-Level Model

In general, it is difficult to solve the bi-level programming problem, as it is a NP-hard problem. Recently, meta-heuristic algorithms (e.g., a genetic algorithm, simulated annealing, etc.) have become the new method of solving the combinational optimal problem. Many scholars solve the bi-level programming problem with the genetic algorithm (GA) (Xu et al., 2004) and the simulated annealing (SA) algorithm (Sahin and Ciric, 1998). For detailed information of the SA, refer to Laarhoven and Aarts (1988). In this study, an improved SA algorithm is proposed to solve the DJTE.

#### 3.2.1 Solution representation and neighbourhood function

A real number encoding scheme is used for solution representation (i.e., $\lambda_{rs} = 1$ represents that the demand multiplier to the prior mean OD demand $d_{rs}$ is 1; $\beta = 0.5$ represents that the multiplier between the mean and variance of the link travel time is 0.5). Under this solution representation, the following neighbourhood function (Equation 25) is used to obtain new solutions $\lambda^1$ and $\beta^1$ from the current solutions $\lambda^0$ and $\beta^0$ respectively. Note that both solutions must satisfy constraints (21) and (22).

$$\lambda^1 = \lambda^0 + \Delta \lambda, \quad \beta^1 = \beta^0 + \Delta \beta$$  \hspace{1cm} (25)

#### 3.2.2 Fitness evaluation

Section 2.4 shows that the upper-level model is a constrained optimisation problem. To simplify the solving of the upper-level problem, the constrained optimisation problem will be first transformed into an unconstrained optimisation problem. As the solutions generated by the
neighbourhood function meet the constraints (21) and (22), only constraint (20) is left to handle. In this study, the penalty function method is used for its simplicity and efficiency. Considering constraint (20), we can update the fitness function as follows:

\[
E = Z_1 + M \left( | \min_{i \in I_e} \hat{x}_a(t) - (1 - \rho) \sum_{i \in I_e} \hat{x}_a(t), 0 | \right) \\
+ M \left( | \min_{i \in I_e} ((1 + \rho) \sum_{i \in I_e} \hat{x}_a(t) - \sum_{i \in l_e} \hat{x}_a(t), 0 | \right)
\]

(26)

where \( M \) is a large positive constant, which can be considered the penalty cost.

3.2.3 Rule to accept new solution and cooling schedule

Let \( E^0 \) and \( E^1 \) be the fitness value of the current and new solution respectively. Set \( \Delta E = E^1 - E^0 \), if \( \Delta E < 0 \), then the new solutions are accepted and considered as the current solution. Otherwise, the new solutions are accepted with the probability \( p = e^{-\Delta E/T} \), where \( T \) is the annealing temperature. This setting allows the SA to accept the worse solution with probability to guarantee that the SA can easily escape from the local optimum. To keep the best solution, we use a memory container to reserve the temporary best solutions, and update the best solutions in time. The cooling schedule is given as

\[
T_{\text{new}} = T_{\text{old}} \times \Delta t
\]

(27)

where \( \Delta t \in (0,1) \) is the temperature decrementing factor. In this cooling schedule, the temperature is kept fixed at each inner loop (Step 3 ~ Step 6 in Section 3.2.5).

3.2.4 Convergence criteria

The convergence criterion of the outer loop (Steps 3 ~ 7 in Section 3.2.5) is that the SA stops when the current temperature is equal to the given final temperature \( T_s \).

The convergence criterion of the inner loop (Step 3 ~ Step 6 in Section 3.2.5) for the standard SA is to meet the given number of iterations, \( N_s \). If \( N_s \) is too small, SA will easily fall into the local optimum, but if \( N_s \) is too large, the SA will spend much time obtaining the optimal solution. However, to guarantee the solution quality, \( N_s \) will usually be given a relatively large number. To overcome the limitations of the standard SA in a lengthy evaluation time, improved convergence criteria are proposed. In this study, the inner loop stops when any one of the following criteria is met: i) the value of the fitness function does not change in the inner loop within \( N_s \) consecutive times; ii) the time of the new solution accepted reaches the given number \( N_i \) or iii) the inner loop iteration reaches the given number \( N_s \).

3.2.5 SA algorithm

The detailed steps of the improved SA used to solve the DJTE are as follows.

**Step 1** Initialise \( T, \Delta t, T_s, N_s, N_i, \) and \( N_s \). Set \( n_x = n_t = n_N = 0 \). Set an initial solution \( \lambda^0 \) and \( \rho^0 \).

**Step 2** Obtain the traffic pattern \( f^0 \) through solving the DTA by the solution algorithm in
Section 3.1 using the given values of $\lambda^0$ and $\beta^0$. Calculate the objective function of the upper level $z^0 = z(\lambda^0, \beta^0, f^0)$. Then, the fitness value of SA, $E^0 = E(\lambda^0, \beta^0, f^0)$ according to Equation (26), can be obtained. Set $E_{best}^0 = E^0, \lambda_{best}^0 = \lambda^0, \beta_{best}^0 = \beta^0, f_{best}^0 = f^0$.

**Step 3** Generate the new neighbouring solutions $\lambda^1$ and $\beta^1$ by Equation (25).

**Step 4** Obtain traffic pattern $f^1$ through solving the DTA by the solution algorithm in Section 3.1 using the given values of $\lambda^1$ and $\beta^1$. Calculate the fitness value of SA $E^1 = E(\lambda^1, \beta^1, f^1)$.

**Step 5** Set $\Delta E = E^1 - E^0$. If $\Delta E \leq 0$ or $e^{-\Delta E/T} \geq \text{random}$, then accept the new solution as the current solution, update $\lambda^0 = \lambda^1, \beta^0 = \beta^1, f^0 = f^1$. If $E_{best}^1 \geq E^1$, then update $E_{best}^1 = E^1, \lambda_{best}^1 = \lambda^1, \beta_{best}^1 = \beta^1, f_{best}^1 = f^1$.

**Step 6** If $(n_s \leq N_s)$ and $(n_t \leq N_t)$, then go to Step 3. Otherwise, go to Step 7.

**Step 7** Set $T = T \times \Delta t$. If $T < T_S$, stop and output the best solution $\lambda_{best}, \beta_{best}, f_{best}$ and the best fitness value $E_{best}^0$ of SA. Otherwise, go to Step 3.

### 4. NUMERICAL EXAMPLE

The small network shown in Figure 1 illustrates the proposed bi-level model formulation and solution algorithms. The network consists of seven links and six nodes and has two simple paths for each OD pair (1,3) and (2,4). The attributes of the seven links in this network are given in Table 2. The journey times on path 1 and traffic flow on link 2 can be observed in this example.

![Figure 1. A small test network](image)

Table 2. Link performance function parameters

<table>
<thead>
<tr>
<th>Link</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_a$ (min)</td>
<td>30</td>
<td>30</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>$C_a$ (pcu/min)</td>
<td>25</td>
<td>25</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>$\omega_a$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>
Due to the symmetric network structure and input data, paths 1 and 4 (paths 2 and 3) will have exactly the same traffic conditions. Hence, only the numerical results on paths 1 and 2 will be investigated. The time of interest in this example is the morning period (6:00-10:00 AM), while the demand departure time interval is taken as one minute. The weighting parameters for the objective function are set as $\gamma_1 = \gamma_2 = 0.5$. The parameter of fitness function is set as $\rho = 0.5$ and the parameter for link travel time covariance in Equation (7) is set as $\theta = 0.5$. The confidence level of path journey time reliability, $\alpha$, is assumed to be 90% in the lower-level DTA problem. The prior traffic demands of each OD pair at the beginning of time interval $k$ is assumed to be calculated through Equation (28)

$$d^{13}(t) = d^{24}(t) = 46.2 \times \exp(-0.006t - 0.8)$$ \hspace{1cm} (28)

Assuming the demand multiplier for the actual demand are: $\lambda^* = [1 \ 1]$, and the multiplier between the mean and variance of the link travel time $\beta^* = 0.25$. With the actual traffic demands defined in Equation (28) and the demand multiplier $\lambda^*$, the observed data (path journey times and link flows) is generated by solving the lower-level DTA model. For the algorithm to solve the DTA, the stopping tolerance $\varepsilon$ is set to 0.00001 and the maximum iteration number $N$ is set to 1000.

For the upper-level model, the upper and lower bounds of demand multiplier $\lambda$ used in constraint (21) are taken as $\lambda^* = [0.5 \ 0.5]$ and $\lambda^* = [1.5 \ 1.5]$ respectively. According to the literature and tests to determine the appropriate parameters for the improved SA approach, the following values are used for the SA parameters in this numerical example: $M = 1000$, $T = 100$, $\Delta t = 0.9$, $T_s = 1$, $N_s = 200$, $N_f = 100$, $N_N = 50$.

To quantify the accuracy of DJTE estimation results, the following root mean squared errors (RMSEs) are used as the performance measures.

$$RMSE_\mu = \sqrt{\frac{1}{N_{path} N_{t_r} \sum \sum \sum (t_p^{rs}(k) - \bar{t}_p^{rs}(k))^2}}$$ \hspace{1cm} (29)

$$RMSE_\sigma = \sqrt{\frac{1}{N_{path} N_{t_r} \sum \sum \sum (\sigma_p^{rs}(k) - \bar{\sigma}_p^{rs}(k))^2}}$$ \hspace{1cm} (30)

$$RMSE_\lambda = \sqrt{\frac{1}{N_{OD} \sum (\lambda_{rs} - \bar{\lambda}_{rs})^2}}$$ \hspace{1cm} (31)

$$RMSE_\beta = |\beta - \bar{\beta}|$$ \hspace{1cm} (32)

where $N_{path}$ is the total number of paths in the network; $N_{t_r}$ is the total number of time intervals; $N_A$ is the total number of links in the network and $N_{OD}$ is the total number of OD pairs in the network. The proposed DJTE model is thus solved, and the results are shown in Table 3.
Table 3. Observed values vs estimated values

<table>
<thead>
<tr>
<th></th>
<th>Observed value</th>
<th>Estimated value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>[1 1]</td>
<td>[0.9091 1.0542]</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.25</td>
<td>0.2517</td>
</tr>
<tr>
<td>RMSE$_\mu$</td>
<td></td>
<td>0.1374</td>
</tr>
<tr>
<td>RMSE$_\sigma$</td>
<td></td>
<td>0.0690</td>
</tr>
<tr>
<td>RMSE$_\lambda$</td>
<td></td>
<td>0.0748</td>
</tr>
<tr>
<td>RMSE$_\beta$</td>
<td></td>
<td>0.0017</td>
</tr>
</tbody>
</table>

In Table 3, it can be seen that the estimated $\lambda$ (0.9091 and 1.0542) and $\beta$ (0.2517) are relatively close to the observed values. By comparing the RMSE$_\mu$ and RMSE$_\sigma$ with the range of mean travel time (28 ~ 38 minutes in Figure 2) and the SD of travel time (7.5 ~ 12 minutes in Figure 3), the proposed model is found to provide a relatively good estimation of the mean and SD, or variance, of travel time. Considering the RMSEs of the parameters, the model gives a more accurate estimation for the multiplier between mean and variance of link travel times ($\beta$).

To obtain a more detailed analysis of the temporal performance of the proposed model in the estimation of path journey times, the observed and estimated mean path journey times (Figure 2), SD of path journey times (Figure 3), effective path journey times (Figure 4) and path inflows (Figure 5) for paths 1 and 2 are plotted and discussed.

Figure 2. Comparison between observed and estimated mean of path journey times

Comparing the difference of the observed and estimated mean journey time of the two paths in Figure 2, the proposed DJTE model performs better in the early morning (from 6:00 AM to 7:00 AM) than in the peak hour period (from 8:00 AM to 9:00 AM). This difference may be due to the increasing variability of demand and the highly dynamic traffic flow propagation processes in the peak hour period. The proposed DJTE model is also found in general to underestimate the mean path journey time for all departure time intervals.

The observed and estimated SDs of the journey times of the two paths show that despite overestimations at particular time intervals (e.g., path 1 at around 7:15 AM), the proposed DJTE model provides a relatively good estimation.
If travellers only consider the mean path journey time when making their route choice, based on the results shown in Figure 2 they will only choose path 2, as the mean path journey times are less than those of path 1. However, both paths have positive inflows (Figure 5), which may be due to the variance of the journey time for path 2 being larger than that of path 1 (Figure 3). Thus, when taking the variance of the path journey time into account, the overall journey time of path 2 may exceed that of path 1. Therefore, in the lower-level model of the proposed DJTE, the reliability of the path journey time is incorporated into the travellers’ route choice when solving the DTA (Equation 6). In this numerical example, a 90% chance of arriving at the destination within the path travel time budget (i.e., $\alpha = 0.9$) is assumed.

The observed and estimated effective path journey times for paths 1 and 2 are shown in Figure 4, and are exactly the same for all departure times as both of the paths are used (Figure 5). The variation of the effective travel time (Figure 4) is similar to that for the mean path journey time (Figure 2) with an increasing trend toward the peak period (8:00 AM ~ 9:00 AM). Similarly, the proposed DJTE model has a larger estimation error in the peak period and in general underestimates the effective path journey time.
Figure 4 clearly illustrates that our DJTE can estimate the effective path journey time for travellers whenever they depart. When the traffic conditions become more congested (increasing demand), the effective path journey time will be greater. When confronted with an increasing path journey time, travellers should depart as earlier as possible to avoid the peak hour period.

Figure 5 shows the observed and estimated path inflows of the two paths. The estimated errors for path 2, which has a larger inflow, are greater than for path 1. When the estimated demand multipliers are equal to \([0.9091 \ 1.0542]\), the demand for the OD pair (1,3) is underestimated and overestimated for the OD pair (2,4). As paths 2 and 3 have a common link 4, travel costs on these two paths are highly correlated. The overestimated demand multiplier for OD pair (2,4) will increase the travel costs on path 3, so the costs on path 2 will also increase. Thus, the path inflow on path 2 will surely be underestimated.

Finally, it should be noted that the accurate estimation of dynamic journey time in a stochastic network is highly related to the selection of weighting factors in Equation (19). In this study, \(\gamma_1 = \gamma_2 = 0.5\) are used, so the mean and SD of the journey time have equal importance. If the mean and SD of the journey time have a bias weighting factor, the estimation results will be inaccurate. These results agree well with those of a static network (Shao et al., 2013), and are not shown here.

5. CONCLUSIONS

A bi-level programming model for DJTE is proposed in this study. To solve the DJTE problem, an SA solution algorithm was modified with a potential global search ability. In the proposed model, the objective of the upper-level problem is a weighted least squares function, minimising the difference of the mean and SD between the observed and estimated path travel times. In the lower-level problem, the reliability-based dynamic traffic assignment model is used to account for travellers’ path choice behaviour under supply uncertainty. A route- or path-based approach using the MSA algorithm is applied to solve the lower-level DTA problem, through which the traffic pattern (e.g., path flows, link flows, path travel times and link travel times) can be obtained. Numerical examples are used to illustrate the application of the proposed DJTE model and the modified SA algorithm.
This study proposes a modified SA algorithm for solving the DJTE problem, which is heuristic in nature. A future research direction is to investigate the potential for using other more efficient algorithms for solving the DJTE problem in realistic networks. Considering demand uncertainty in a road network and incorporating other types of link travel time functions into the reliability-based DTA model will also be beneficial.

ACKNOWLEDGEMENTS

The work described in this paper was jointly supported by grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project Nos. PolyU 5242/12E and 152074/14E), the Research Institute for Sustainable Urban Development (RISUD) of the Hong Kong Polytechnic University (Project Nos. 1-ZVBZ and 1-ZVEW), the Hong Kong Scholars Program (Project No. G-YZ21), and the National Natural Science Foundation of China (Nos. 71371001, 71171013, and 71471013).

APPENDIX A – COMPUTATION OF MEAN AND SD OF THE PATH JOURNEY TIMES

Assume a path \( p \) between OD pair \((r,s)\) consists of the following sets of links \( \{a_1,a_2,\cdots,a_m\} \) and nodes \((r,1,2,\cdots,m-1,s)\). Obviously,

\[
\begin{align*}
\mu_p^r(k) &= t_{a_1}(k) \\
\sigma^r_p(k) &= \sigma_{a_1}(k) \\
\mu_p^r(t) &= t_{a_1}(t) + t_{a_2}(t + t_{a_1}(t)) \\
(\sigma^r_p(t))^2 &= (\sigma_{a_1}(t))^2 + (\sigma_{a_2}(t + t_{a_1}(t)))^2 + 2 \text{cov}(t_{a_1}(t), t_{a_2}(t + t_{a_1}(t))) \\
&\cdots \\
\end{align*}
\]

We can then express the mean of the path travel time on path \( p \in P_{rs} \) of pair \((r,s)\) at the end of time interval \( k \in T_d \), \( \mu^r_p(t) \), by recursively applying the mean of the link travel time.

\[
\begin{align*}
\mu^r_p(t) &= t_{a_1}(t) + t_{a_2}(t + t_{a_1}(t)) + \cdots + t_{a_n}(t + t_{a_{n-1}}(t) + t_{a_n}(t)) + \cdots + t_{a_m}(t + t_{a_{m-1}}(t) + \cdots + t_{a_m}(t)) \\
&= \sum_{i=1}^{m} t_{a_i}(t + t_{a_{i-1}}(t) + \cdots + t_{a_m}(t)) \\
(\sigma^r_p(t))^2 &= (\sigma_{a_1}(t))^2 + (\sigma_{a_2}(t + t_{a_1}(t)))^2 + \cdots + (\sigma_{a_n}(t + t_{a_{n-1}}(t) + \cdots + t_{a_m}(t)))^2 \\
&\quad + 2 \sum_{1 \leq i \leq m} \text{cov}(t_{a_i}(t + t_{a_{i-1}}(t) + \cdots + t_{a_m}(t)), t_{a_i}(t + t_{a_{i-1}}(t) + \cdots + t_{a_m}(t))) \\
&= (\sigma_{a_1}(t))^2 + (\sigma_{a_2}(t + t_{a_1}(t)))^2 + \cdots + (\sigma_{a_n}(t + t_{a_{n-1}}(t) + \cdots + t_{a_m}(t)))^2 \\
&\quad + 2 \sum_{1 \leq i \leq m} \text{cov}(t_{a_i}, t_{a_i}) \\
\end{align*}
\]

where \( t_{a_i} = t_{a_i}(t), t_{a_{i+1}} = t_{a_1}(t + t_{a_1}(t)) \) for short. We can rewrite Equations (A1) and (A2) as
\[ t_p^{rs}(t) = \sum_{a \text{ on path } p, t \in [k \Delta t], k \in T_f} \sum_{l} t_a(l) \delta_{apk}^{rs}(l) \tag{A3} \]
\[ (\sigma_p^{rs}(t))^2 = (\sum_{a \text{ on path } p, t \in [k \Delta t], k \in T_f} \sigma_a(l) \delta_{apk}^{rs}(l))^2 + 2 \sum_{l \in [s \Delta t] \land m} \text{ cov}(t_{a}, t_{a'}) \tag{A4} \]

where \( \delta_{apk}^{rs}(l) \) is equal to 1, if the flow on path \( p \) of OD pair \((r,s)\) enters the network at interval \( k \) and arrives at link \( a \) at interval \( l \); otherwise, 0. The detailed information is as follows:

\[ \delta_{apk}^{rs}(l) = \begin{cases} 
1 & \quad \text{If } k + t_{a_1} + t_{a_2} + \cdots + t_{a_{l-1}} = l \\
0 & \quad \text{Otherwise.} 
\end{cases} \]

And for any link \( a \) on path \( p \), clearly

\[ \sum_{l \in [k \Delta t], k \in T_f} \delta_{apk}^{rs}(l) = 1 \quad \forall p \in P_{rs}, rs \in R, k \in T_d \]

REFERENCES


