An Activity-Based Model for the Estimation of Origin-Destination Matrix on Road Networks

Lian Qun OUYANG, Di HUANG

Abstract: Travel demand forecasting in terms of origin-destination (*O-D*) matrix estimation on transportation networks is an important topic in transportation research. The motivation of each trip is the user's desire of conducting a compulsory or noncompulsory activity. User activity choices should not be ignored for the trip demand forecasting; nevertheless, there are scare approaches of *O-D* matrix estimation incorporating users' activity behavior in the literature. This study develops an activity-based bi-level model for the estimation of *O-D* matrix on road networks using the activity-based approach where road users' activity and travel choices are integrated scheduled. In the estimation, the parameter of the user activity and travel choice model is simultaneously calibrated. A heuristic algorithm is explored to solve the model. A numerical example is provided. The example results illustrate that the developed model and heuristic algorithm are applicable and efficient tools to deal with the problem of *O-D* matrix estimation.

Keywords: the activity-based approach; network model; origin-destination matrix

1 INTRODUCTION

There appeared a mass of literature on the discussion of the problem of origin-destination (*O-D*) matrix estimation on transportation networks in the past decades due to the important role of the accuracy forecasting of trip demands in transportation management. Various estimation approaches with the use of the observation data on some selected links of the transportation network have been developed in the literature [1-11].

In the earlier estimation models and methods, traffic congestions on the transportation network did not consider where users were supposed to make travels independently [1]. Later, traffic congestions resulting from the interactions among users' travels on the network were studied [2-5]. In those studies, usually a network equilibrium model was formulated to describe the user aggregation travel behavior. Some attempts have been made to simultaneously estimate the *O-D* matrix on transportation networks and calibrate the parameters employed in the equilibrium model of user travel choices [6-7].

In recent studies, the fuzzy method has been adopted for the estimation of the *O-D* matrix [8], especially for the estimation while the observation data is incomplete or imprecise [9]. Some researchers [10-11] recently discussed how to apply the new technologies such as internet routing data or global position system (GPS) data to estimate the *O-D* matrix.

Most previous *O-D* estimation models including the above were the trip-based approach. The effect of road users' activity behavior on *O-D* matrices was ignored. That ignorance possibly leads to the misrepresentation of users' travel behavior [12] and the biased assessment of the effects of transport policies [13]. The underlying generation of trips on the transportation network is the need of the trip makers to conduct compulsory or non-compulsory activities at different locations. For example, the trip from home to the supermarket is generated by the need of shopping. The destination of the trip is the target location for the activity conduction while the origin of the trip is the stay location of the trip maker before traveling. If there is no desire of activity conductions, no trip demand is generated. Therefore, for the achievement of the accuracy trip demand forecasting, the analysis of user activity behavior should be incorporated into the estimation of *O-D* matrix.

Recently, the activity-based approach has been widely adopted to develop activity-based transport models for the better understanding of user travel behavior [14-22]. In the developed models, not only users' trips on the transportation network, but also users' activities, trip chain, activity chain, and the coordination of activities and trips were discussed. It is noticed that most of the early activity-based transport models are the simulation models because it is difficult to clarify the complex relationships between activities and travels of users by mathematical expressions [15-17].

With the improvement of research on the activity-based approach, a number of mathematical activity-based transport models have been proposed [14, 18-22] in the past recent years. Huang and Lam [14] discussed the development of an analytical model of scheduling users' time-dependent activities and travels. Joh et al. [18] conducted a research study to develop a model for the activity rescheduling problem. Various supernetwork approaches have been explored to deal with the activity scheduling problems [19-22]. With the use of those supernetwork approaches, the daily time-dependent trip demands on transportation networks could be endogenously generated.

However, little attention has been paid to employ the activity-based approach in the estimation of *O-D* matrix in the previous studies, although there exist strong interactions between activities and trip demands. This study investigates the problem of *O-D* matrix estimation on road networks with the consideration of road users' activity choices. To deal with that problem, an activity-based bi-level model is proposed. In the model, the history *O-D* matrix and the vehicular count information on some selected road links are the base data for the estimation. In the estimation, the impact of user activity behavior on the *O-D* matrix is examined and the parameter of user activity and travel choice model is calibrated.

The organization of the paper is as follows. Section 2 discusses the problem of user activity and travel choices. Section 3 presents the formulation of the activity-based bi-level model. In Section 4, a heuristic algorithm is explored to solve the formulated model. Section 5 presents the applications of the method developed in this study on a

sample transportation network. Section 6 gives a final summary together with the research directions in future.

2 USER ACTIVITY AND TRAVEL CHOICES
2.1 User Activity and Travel Choice Model

In this section, the road users' behavior of activities travels on transportation networks is discussed. Users are supposed to make activity and travel decisions with the maximum expected utility principle.

Let N and J denote the node set on the transportation network and the activity set including all considered compulsory activities and noncompulsory activities, respectively. Let C_{rs} denote the travel time of the trip from origin node r to destination node s, where $r \in N$ and $s \in N$. Let μ_s^j represent the utility of a user conducting activity j at activity location node s, where $j \in J$. Let α be the equivalent travel time of unit utility. Let ε_{sj} represent the random perception errors of the utility, and it is assumed to follow the independent and identical Gumble distribution.

The total expected utility V_{rs}^j of a user making a trip from node r to node s and then conducting activity j at node s, is formulated as the following equation.

$$V_{rs}^j = -C_{rs} + \alpha\mu_s^j + \varepsilon_{sj}, \quad \forall r \in N, s \in N, j \in J \quad (1)$$

The probability, p_{rs}^j, of a road user at origin node r traveling to destination node s and then conducting activity j at node s, is calculated by

$$p_{rs}^j = \frac{\exp[(-C_{rs} + \alpha\mu_s^j)\lambda]}{\sum_{s,j}\exp[(-C_{rs} + \alpha\mu_s^j)\lambda]}, \quad \forall r, s \in N, j \in J. \quad (2)$$

Let d_r denote the total number of road users at origin node r. The user activity and travel choice model which generates the demand of users, d_{rs}^j, who are choosing traveling from node r to node s and then conducting activity j at node s, can be formulated as

$$d_{rs}^j = d_r \frac{\exp[(-C_{rs} + \alpha\mu_s^j)\lambda]}{\sum_{s,j}\exp[(-C_{rs} + \alpha\mu_s^j)\lambda]}, \quad \forall r, s \in N, j \in J. \quad (3)$$

In the model of Eq. (3), λ is a parameter needing to be calibrated when the history value of λ is unable to properly capture users' updated activity and travel behavior.

There may be more than one activity location nodes on the transportation network feasible for a specific activity conduction. It is also possible that there is more than one alternative activity at an activity location node. Let J_s denote the set of alternative activities at activity location node s, where $J_s \subseteq J$. Let V_s' denote the expected utility

of a user conducting any available activity at activity location node s, then,

$$V_s' = \ln \sum_{j \in J_s} \exp(\alpha\mu_s^j), \quad \forall s \in N. \quad (4)$$

Let R denote the set of origin nodes on the transportation network, where $R \subseteq N$ and $\forall r \in R$. The probability, p_{rs}, of a user choosing the trip from node r to node s, is

$$p_{rs} = \frac{\exp(-C_{rs} + V_s')}{\sum_{l \in R}\exp(-C_{rl} + V_l')}, \quad \forall r, s \in N. \quad (5)$$

2.2 The Activity-based Stochastic User Equilibrium Model

In this section, an activity-based stochastic user equilibrium model (SUE) is proposed to deal with the users' activity and travel choice problem on the road network.

Let p and a, respectively, denote a path and a link of the studied road network. Let P_{rs} denote the set of all feasible paths from node r to node s on the network. Let $f_{rs,p}^j$ represent the demand traveling from node r to node s, choosing path p and then conducting activity j at node s. Let x_a^j represent the demand experiencing link a for conducting activity j purpose. Let $\delta_{a,p}^{rs}$ be equal to 1 if link a is experienced by path p where path p is a feasible path from node r to node s, and 0 otherwise.

The constraints of the above discussed user activity and travel choice problem on the road network are summarized as follows:

$$\sum_p f_{rs,p}^j = d_{rs}^j, \quad \forall r, s \in N, p \in P_{rs}, j \in J \quad (6a)$$

$$x_a^j = \sum_r \sum_s \sum_p f_{rs,p}^j \delta_{a,p}^{rs}, \quad \forall r, s \in N, p \in P_{rs}, j \in J \quad (6b)$$

$$\sum_j \sum_s d_{rs}^j = d_r, \quad \forall r \in N, s \in N, j \in J \quad (6c)$$

$$f_{rs,p}^j \geq 0, \quad \forall p \in P_{rs}, r, s \in N, j \in J \quad (6d)$$

$$d_{rs}^j \geq 0, \quad \forall r, s \in N, j \in J. \quad (6e)$$

According to the maximum expected utility principle, the road users' choices of activities and travels on the network result in a SUE state, which can be formulated as follows,

$$\left(T_{rs}^j - T_r^*\right)d_{rs}^j = 0, \quad \forall r, s \in N, j \in J \quad (7a)$$

$$T_{rs}^j \geq T_r^*, \quad \forall r, s \in N, j \in J \quad (7b)$$

$$d_{rs}^j \geq 0, \quad \forall r, s \in N, j \in J \quad (7c)$$

$$T_r^* = \min_{s,j} T_{rs}^j = \min_{s,j}\left(C_{rs} - \alpha\mu_s^j\right), \quad \forall r \in N \quad (7d)$$

where T_{rs}^j is the total expected disutility of traveling from node r to node s and then conducting activity j at node s

and it is the systematic (or deterministic) component of total utility of V_{rs}^j in Eq. (1). Eq. (7a)-(7d) imply that the expected disutility of activity and travel choices of each road user that departs from the same origin node, is equal and minimized at the equilibrium state.

Let $C_{rs,p}^{j*}$ denote the expected travel time of users traveling from node r to node s, choosing path p and then conducting activity j at node s at the equilibrium state. Let d_{rs}^{j*} represent the corresponding d_{rs}^j at the equilibrium state. Let $f_{rs,p}^{j*}$ represent the corresponding $f_{rs,p}^j$ at the equilibrium state.

Theorem 1. The traffic flow on the road network satisfying constraints (6a)-(6e) reaches an activity-based SUE state if and only if it is the solution to the below variational inequality (VI) problem.

$$\sum_j \sum_r \sum_s \sum_p C_{rs,p}^{j*} \left(f_{rs,p}^j - f_{rs,p}^{j*} \right)$$
$$+ \sum_j \sum_r \sum_s \left(\frac{1}{\lambda} \ln \frac{d_{rs}^{j*}}{d_r} - \alpha \mu_s^j \right) \left(d_{rs}^j - d_{rs}^{j*} \right) \geq 0, \tag{8}$$

Proof. The VI problem of Eq. (8) is considered as a compound of two parts. Part 1 of Eq. (8) is the below equation,

$$\sum_j \sum_r \sum_s \frac{1}{\lambda} \ln \frac{d_{rs}^{j*}}{d_r} \left(d_{rs}^j - d_{rs}^{j*} \right) \geq 0, \tag{9}$$

subject to constraints (6c) and (6e).

It is not difficult to prove that Part 1 of Eq. (8) is equivalent to the following programming problem [23]:

$$\min \sum_j \sum_r \sum_s \frac{1}{\lambda} d_{rs}^j \left(\ln d_{rs}^j - 1 \right), \tag{10}$$

subject to constraints (6c) and (6e). Eq. (10) is equivalent to Eq. (3) which describes the road user activity and travel behavior.

Part 2 of Eq. (8) is,

$$\sum_j \sum_r \sum_s \sum_p C_{rs,p}^{j*} \left(f_{rs,p}^j - f_{rs,p}^{j*} \right)$$
$$- \sum_j \sum_r \sum_s \alpha \mu_s^j \left(d_{rs}^j - d_{rs}^{j*} \right) \geq 0, \tag{11}$$

subject to constraints (6a)-(6d).

In view of Eq. (6a), Part 2 of Eq. (8) could be rewritten as,

$$\sum_j \sum_r \sum_s \left(C_{rs,p}^{j*} - \alpha \mu_s^j \right) \left(d_{rs}^j - d_{rs}^{j*} \right) \geq 0 . \tag{12}$$

The formulation of Eq. (12) is the equivalent VI problem of Eq. (7a)-(7d). In view of the formulations of Eq. (10) and Eq. (12), Theorem 1 holds.

3 THE ACTIVITY-BASED BI-LEVEL MODEL

An activity-based bi-level model is developed in this section. There are two objectives of the model development in this study. One is that the **O-D** matrices on road networks with the consideration of user activity behavior are estimated with the use of the observation data. The other one is that the parameter λ of the user activity and travel choice model of Eq. (3) is calibrated to keep up with the updated user activity and travel behavior.

3.1 The Upper-level Problem

The upper-level problem is the estimation objective. Assume that the history **O-D** matrix and the history value of parameter λ are given. The vehicular counts on some selected road links of the transportation network are available through the observations [24].

For simplifying the presentation, the vectors of variables are used in the discussions of this section. Let **d** denote the updated **O-D** demands after the estimation. Let $\overline{d}$ denote the history of **O-D** demands. Let **x** represent the estimated flows on the observation road links. Let $\overline{x}$ represent the observed flows on the observation road links. Let $\boldsymbol{P}(\lambda)$ be the proportions of link flows on each **O-D** pair flow. $\boldsymbol{A}^{-1}$ is the predetermined weighting of the **O-D** demand estimation errors on the estimation results while $\boldsymbol{B}^{-1}$ is the predetermined weighting of the link flow estimation errors on the estimation results. Note that the sum of $\boldsymbol{A}^{-1}$ and $\boldsymbol{B}^{-1}$ is 1. Readers may refer to the study of [31] for the value settings of $\boldsymbol{A}^{-1}$ and $\boldsymbol{B}^{-1}$.

In the **O-D** matrix estimation problem, it is easy to understand that the smaller value gap between the estimated demands/flows and the history demands/ observed flows, is a better estimation. Therefore, the upper-level problem is formulated as the following equations,

$$\min Z(\boldsymbol{d}, \boldsymbol{x}, \lambda) = \left(\overline{\boldsymbol{d}} - \boldsymbol{d}(\lambda) \right)^{\mathrm{T}} \boldsymbol{A}^{-1} \left(\overline{\boldsymbol{d}} - \boldsymbol{d}(\lambda) \right) +$$
$$+ \left(\overline{\boldsymbol{x}} - \boldsymbol{x}(\boldsymbol{d}, \lambda) \right)^{\mathrm{T}} \boldsymbol{B}^{-1} \left(\overline{\boldsymbol{x}} - \boldsymbol{x}(\boldsymbol{d}, \lambda) \right) \tag{13a}$$

subject to

$$\boldsymbol{x}(\boldsymbol{d}, \lambda) = \boldsymbol{P}(\lambda)\boldsymbol{d}. \tag{13b}$$

Eq. (13a) is the objective function of the estimation problem while Eq. (13b) is the constraint equation which expresses the relationship between link flows and **O-D** demands on the transportation network. The constrained generalized least square (GLS) method [25] which was widely used in the previous relevant estimation studies, could be adopted to solve the upper-level problem of Eq. (13a) and Eq. (13b).

3.2 The Lower-Level Problem

The lower-level problem is the user activity and travel behavior in responding to the updated **O-D** matrix and the updated value of λ. According to the discussions in

Section 2, it is an activity-based SUE problem and it is formulated as follows,

$$\sum_j \sum_r \sum_s \sum_p C_{rs,p}^{j*} \left(f_{rs,p}^j - f_{rs,p}^{j*} \right)$$

$$+ \sum_j \sum_r \sum_s \left(\frac{1}{\lambda} \ln \frac{d_{rs}^{j*}}{d_r} - \alpha \mu_s^j \right) \left(d_{rs}^j - d_{rs}^{j*} \right) \geq 0, \quad (13c)$$

$$\sum_p f_{rs,p}^j = d_{rs}^j, \qquad \forall r,s \in N, p \in P_{rs}, j \in J \quad (13d)$$

$$x_a^j = \sum_r \sum_s \sum_p f_{rs,p}^j \delta_{a,p}^{rs}, \quad \forall r,s \in N, p \in P_{rs}, j \in J \quad (13e)$$

$$\sum_j \sum_s d_{rs}^j = d_r, \qquad \forall r \in N, s \in N, j \in J \quad (13f)$$

$$f_{rs,p}^j \geq 0, \qquad \forall p \in P_{rs}, r,s \in N, j \in J \quad (13g)$$

$$d_{rs}^j \geq 0, \qquad \forall r,s \in N, j \in J. \quad (13h)$$

4 HEURISTIC ALGORITHM

A heuristic algorithm is explored to solve the developed activity-based bi-level model of Eq. (13a)-(13h). In the iterations of the heuristic algorithm, the Bureau of Public Roads (BPR) functions are used to calculate the travel time $t_a(x_a)$ of users on each road link.

Let $t_a(0)$ denote the travel time on link a while the link flow is zero. Let x_a be the flow on link a. Let s_a be the link capacity. The below BPR function is used:

$$t_a(x_a) = t_a(0) \left(1 + 0.15 \left(\frac{x_a}{s_a} \right)^4 \right). \quad (14)$$

The heuristic algorithm is designed as follows:

Step 1: Set the iteration number and let $n = 1$. Set the value of the history O-D matrix as $d^{(n)}$. Set the history value of λ as $\lambda^{(n)}$. Calculate the initial link flows $x^{(n)}$ and initial link flow proportions $P^{(n)}$ by solving the lower-level problem with the given $d^{(n)}$ and $\lambda^{(n)}$.

Step 2: Use the GLS method [25] to obtain the estimated O-D matrix d^*. Calculate the optimum iteration step of σ which is the solution to the problem of Eq. (13a) with the values of d^*, $x^{(n)}$ and $\lambda^{(n)}$, i.e., $\min Z(d^*, x^{(n)}, \lambda^{(n)}, \sigma)$. Update the O-D matrix by the equation of $d^{(n+1)} = d^{(n)} + \sigma(d^* - d^{(n)})$.

Step 3: Update link flows $x^{(n+1)}$ by solving the lower-level problem according to the updated $d^{(n+1)}$ and $\lambda^{(n)}$. Note that $x^{(n+1)}$ relates with $d^{(n+1)}$ and $\lambda^{(n)}$.

Step 4: Use the constrained GLS method to obtain the calibrated value of λ^*. Calculate the optimum iteration step of γ which is the solution to the problem of Eq. (13a) with the values of $d^{(n+1)}$, $x^{(n+1)}$ and $\lambda^{(n)}$, i.e., $\min Z(d^{(n+1)}, x^{(n+1)}, \lambda^{(n)}, \gamma)$. Update the value of the parameter by the equation of $\lambda^{(n+1)} = \lambda^{(n)} + \gamma(\lambda^* - \lambda^{(n)})$.

Step 5: Update the link flows $x^{(n+1)}$ by solving the lower-level problem according to the updated $d^{(n+1)}$ and and $\lambda^{(n+1)}$.

Step 6: Check whether the convergence condition is satisfied or not. If satisfied, stop the iterations; otherwise, let $n = n+1$ and go to Step 2.

The convergence condition in Step 6 could be the below equation where the maximum relative change in the elements of the estimated O-D matrix and the calibrated parameter value at successive iterations should be less than a small positive number:

$$\max_i \left\{ \left(\left| d_i^{(n+1)} - d_i^{(n)} \right| / d_i^{(n)} \right), \left(\left| \lambda^{(n+1)} - \lambda^n \right| / \lambda^n \right) \right\} \leq \eta. \quad (15)$$

In Eq. (15), η is usually set as a positive number close to zero.

5 NUMERICAL EXAMPLE

In this section, a numerical example is provided to illustrate the application of the developed activity-based bi-level model and the explored heuristic algorithm in solving the problem of estimating O-D matrix on road networks. The convergence of the heuristic algorithm is examined. The accuracy of the estimation results with the application of the methods in this study is discussed.
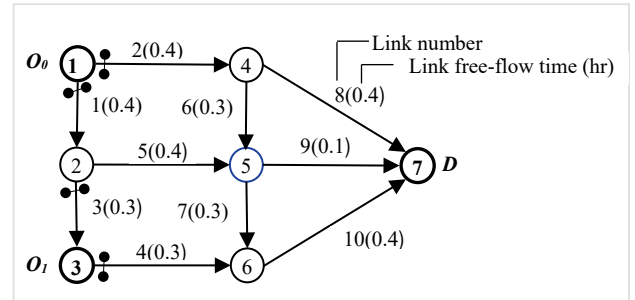

Figure 1 An example road network

Fig. 1 gives the considered road network. It is a network of 7 nodes and 10 road links. In this example, the unit of the link travel times is in hours (h). The unit of the link flows and link capacities is in number of vehicles per hour (veh/h). The free-flow travel time of each road link of the network is shown in Fig. 1. The capacity of each road link of the network is assumed to be 4000 veh/h. Link 1, Link2, Link 3, and Link 4 are the selected links where the number of passed vehicles is counted by the observation device [26, 27]. Node 1 and Node 3 are two residential areas, Node 7 is a working zone and there is a restaurant at Node 5 in the example network.

Suppose that, every morning, users at Node 1 and Node 3 (i.e. the origin nodes) need to travel to Node 7 (i.e. the destination node and activity location node) for work activity, and on that journey before the arrival to work users may choose whether to have breakfast at Node 5 (i.e. the activity location node) or not. The utilities of the activities are assumed with fixed values of 60 for work at Node 7 and of 15 for having breakfast at Node 5, respectively.

Road users' trips from Node 1 to Node 7 are defined as **O-D** pair 0, and road users' trips from Node 3 to Node 7 are defined as **O-D** pair 1. The total number of road users at the original node of Node 1 d_1 is 2000 veh/h and that at the other original node of Node 3 d_2 is 3000 veh/h. The activity chain choice of having breakfast at Node 5 before work is defined as AC0. The activity chain choice of non-breakfast at Node 5 before work is defined as AC1.

The actual value of parameter λ for the individual activity and travel choice model of Eq. (3) is assumed as 0.1. The value of parameter α in Eq. (1) is set as 0.1. Tab. 1 gives the actual and history **O-D** matrices. In addition, the actual and counted link flows on the selected

observation road links are presented in Tab. 1. Note that the history value of λ is 0.05.

Through the comparisons between the actual and the history values of λ and the comparisons between the actual and the history **O-D** matrices in Tab. 1, it is easy to find that the history λ and the history **O-D** matrix are greatly biased. There is a need to estimate the **O-D** matrix and calibrate the value of λ which is used to model user activity and travel behavior. The surveyed vehicular counts information is used to deal with the estimation and calibration problem. Because the history data is rather inaccurate and the observation counts seem to be much more credible in this example, we set $A^{-1} = 0$ and $B^{-1} = 1$.

Table 1 The actual/history **O-D** matrix (vehicles) and observed link flows on the selected links (vehicles per hour)

| **O-D**\Activity choices	Actual **O-D** matrix			History **O-D** matrix		
	AC0 (stop at Node 5)	AC1	Total	AC0	AC1	Total
1-7(**O-D** 0)	1291	709	2000	700	800	1500
3-7(**O-D** 1)	1936	1064	3000	1000	1600	2600
Counted link flow\Observed link	1	2		3	4	
Counted flow of AC0	457	834		1223	714	
Counted flow of AC1	0	709		1	1063	

5.1 Analysis of the Results

Firstly, the convergence of the explored heuristic algorithm is investigated. The value of the parameter of the convergence condition of Eq. (15) is set as $\eta = 0.0001$. Three different initial points, i.e. history **O-D** matrices in view of the discussions in Section 4, are tested in the investigation as shown in Fig. 2.

From the illustration of Fig. 2, under various history **O-D** matrices, the estimation and calibration results could still reach convergence with only a few iterations. Those results demonstrate that the explored heuristic algorithm in Section 4 is a convergent and efficient solution algorithm for solving the activity-based bi-level problem considered in this study.

Secondly, the estimation and calibration results of the **O-D** matrix and link flows with different activity choices are discussed in Tab. 2. It can be seen that the estimation results of **O-D** matrix with activity choices are close to the

actual **O-D** matrix as given in Tab. 1. On the observation road links, the estimated link flows are also close to the counts as given in Tab. 1. In addition, the calibration result of the value of λ is 0.1007, which is also close to the actual value of 0.1. Therefore, the estimation and calibration results are qualified.
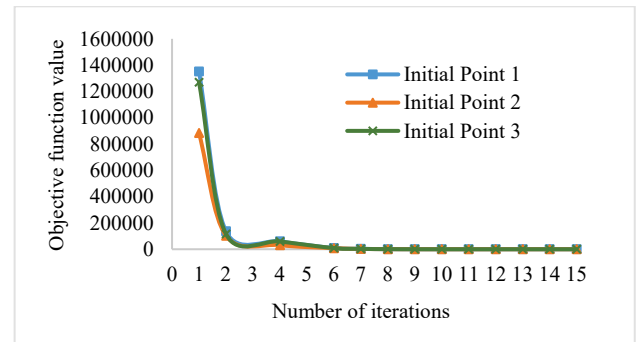
Figure 2 Convergence test of the explored heuristic algorithm

Table 2 Estimated **O-D** matrix and link flows by different activity choices (vehicles per hour)

O-D\Activity choices	AC0 (stop at Node 5)			AC1			Total			
1-7(**O-D** 0)	1295			710			2005			
3-7(**O-D** 1)	1938			1063			3001			
Activity choices\Link	1	2	3	4	5	6	7	8	9	10
AC0	461	834	1222	716	1683	834	716	0	3234	0
AC1	0	710	1	1062	1	0	0	710	1	1062

Except that the accuracy of the estimation and calibration results is guaranteed, the results with the application of methods in this study also provide more information of user activity and travel choices. For example, from Tab. 2, not only the **O-D** trip demands are estimated but also the **O-D** trip demands with different activity purposes are obtained. That may be difficult to be achieved by using the conventional **O-D** matrix estimation models. Furthermore, road users' travel behavior of

temporary stopping at Node 5 before work in this example, could be examined but that behavior is probably ignored in the conventional **O-D** matrix estimation approaches.

Road users' activity choices are constrained by both the transportation network and the activity location distributions. In turn, the demands for various activities make important impacts on the development of those activity areas for matching demands. Therefore, the estimation of **O-D** matrix with activity purposes, probably

not only provide more information for improving management and planning on traffic policies and transportation infrastructure, but also could support a way to how to optimize the land use to decrease the urban traffic congestion.

5.2 Sensitivity Tests

Thirdly, the effects of activity utility functions and the potential number of road users over the network on users' activity and travel behavior are examined.

Suppose that the service quality of the restaurant at Node 5 drops in standard, and the utility of having breakfast at Node 5 declines from 15 to 5. The effects of such a change on users' behavior are discussed in Tab. 3. As shown, while the utility of having breakfast at Node 5 decreases, some road users who had breakfast at Node 5 before, choosing AC0, cancel the eating at Node 5 and switch to the activity chain AC1.

Suppose that the number of the original road users at Node 1 and Node 3 grows from 2000 veh/h to 6000 veh/h, and from 3000 veh/h to 8000 veh/h, respectively. The increased number of original population may lead to more serious traffic congestions on the network. The impact of that increase on users' activity and travel behavior is presented in Tab. 3. It is shown that when the traffic congestion becomes worse, more road users prefer to choose AC1, with no stop at Node 5 for eating.

Those results shown in Tab. 3 imply that the developed model and explored heuristic algorithm in this study are applicable to capture the effects of the environmental changes such as land use development policies which influence users' activity choice decisions on the **O-D** matrix. However, the conventional approaches appear to fail to incorporate such kinds of effects in their **O-D** matrix estimation.

Finally, the sensitivity test of the parameter λ for the individual activity and travel choice model of Eq. (3), is conducted. For simplicity, let the set {{a; b}, {c; d}} denote the activity-based trip demands of the two **O-D** pairs, where a, b, c, d are positive integers. For example, in Tab. 4, {{1298; 717}, {1943; 1061}} representing the **O-D** flow from Node 1 to Node 7 for AC0 and AC1 are 1298 and 717, respectively, while the **O-D** flows from Node 3 to Node 7 for AC0 and AC1 are 1943 and 1061, respectively. The basic data are presented in Tab. 1 and the weighting factors $U^{-1} = 0.1$ and $W^{-1} = 0.9$ are used. The **O-D** matrix is estimated in two scenarios. Scenario 1 uses the proposed bi-level model with parameter λ calibration, while scenario 2 is without λ calibration where λ is fixed as 0.05. The results under those two scenarios are compared in Tab. 4.

It is found that the estimation and calibration results in scenario 1 with the parameter λ calibration are quite close to the actual **O-D** matrix and the actual value of λ. In contrast, the estimation errors increase in scenario 2 without the parameter λ calibration. In other words, the accuracy of the estimation results in scenario 2 decreases. Therefore, it is necessary to take into account the effects of the activity choices in the **O-D** matrix estimation problem, and update the parameter that captures user activity and travel behavior. Otherwise, the ignorance of those effects and the parameter calibration may lead to biased **O-D** matrix estimation results.

Table 3 Effects of changes in environment (vehicles per hour)

		O-D 1-7 for AC0	**O-D** 1-7 for AC1	**O-D** 3-7 for AC0	**O-D** 3-7 for AC1
Uncongested and eating utility = 15	Estimated demand	1295	710	1938	1063
	Percentage (%)	64,6	35,4	64,6	35,4
Uncongested and eating utility = 5	Estimated demand	1103	900	1655	1350
	Percentage (%)	55,2	44,8	55,2	44,8
Congested and eating utility = 5	Estimated demand	3268	2736	4357	3638
	Percentage (%)	54,4	45,6	54,4	45,6

Table 4 Results of the scenario 1 and scenario 2, i.e. with/without calibrating the parameter of the activity and travel choice model of Eq. (3)

	Estimated λ	Absolute error of λ	Estimated **O-D** matrix	Absolute error of **O-D** matrix
With the parameter calibration	0,1009	0,0009	{{1298; 717}, {1943; 1061}}	{{7; 8}, {7; 3}}
Without the parameter calibration	0,0500	0,0500	{{1140; 831}, {1739; 1250}}	{{151; 122}, {197; 186}}

The actual data are given as: λ=0.1 and **O-D** matrix= {{1291; 709}, {1936; 1064}}.

6 CONCLUSIONS

The activity-based approach could provide improved understanding of trip demands; however, there is scare literature on the study of incorporating activity choices in **O-D** matrices estimation problems. This is mainly due to the difficulty in collecting adequate data on users' activity behavior. The activity-based bi-level programming estimation model proposed in this study, overcomes the shortcomings of large amounts of data collection.

The proposed model examines the effect of road users' activity pattern on the **O-D** matrix by adopting the activity-based SUE problem as the lower level. The **O-D** matrix of the whole road network is estimated by vehicular counts on some links, while the parameter of the activity and travel choice model has been calibrated simultaneously, in the upper level. Thus, the proposed model can reduce the requirement of a large amount of data to represent all road users' activity choice behavior.

The numerical example results illustrate that the developed estimation model and the explored heuristic algorithm are applicable and efficient. The estimation model not only provides the **O-D** matrix for various activity purposes but also captures some effects of environmental changes. However, conventional estimation models probably fail to do this. More improved information is also offered with the applications for the strategic transport infrastructure projects, policy requirements and land use plans.

This study is the first step of developing activity-based approaches for the **O-D** matrix estimations. The future work is expected to produce formulations to describe users' more comprehensive activity choice behavior. In this study, for simplicity, the network capacity constraints are not considered. In a further study, we can adopt the research methodologies presented in our previous study of [21] to consider the capacity constraint problem. This study focuses on the impact of activity patterns on trip demands in static networks. However, in reality, the proposed activity plays an important role in users' departure time choices. The model should therefore be extended to dynamic networks, together with the consideration of activity durations.

Acknowledgements

7 REFERENCES

[1] Fisk, C. S. (1989). Trip Matrix Estimation from Link Traffic Counts: The Congested Network Case. *Transportation Research B*, 23, 331-336.
https://doi.org/10.1016/0191-2615(89)90009-X

[2] Bell, M. G. H. (1991). The Estimation of Origin-destination Matrices by Constrained Generalized Least Squares. *Transportation Research B*, 25, 13-22.
https://doi.org/10.1016/0191-2615(91)90010-G

[3] Florian, M. & Chen, Y. (1995). A Coordinate Descent Method for the Bi-Level O-D Matrix Adjustment Problem. *International Transactions on Operations Research*, 2, 165-179.

[4] Yang, H., Sasaki, T., Iida, Y., & Asakura, Y. (1992). Estimation of Origin-Destination Matrices from Link Traffic Counts on Congested Networks. *Transportation Research B*, 26, 417-434. https://doi.org/10.1016/0191-2615(92)90008-K

[5] Maher, M. J., Zhang, X., & Vliet, D. V. (2001). A Bi-level Programming Approach for Trip Matrix Estimation and Traffic Count Problems with Stochastic User Equilibrium Link Flows. *Transportation Research B*, 35, 23-40.
https://doi.org/10.1016/S0191-2615(00)00017-5

[6] Yang, H., Meng, Q., & Bell, M. G. H. (2001). Simultaneous Estimation of the Origin-destination Matrices and Travel-cost Coefficient for Congested Networks in a Stochastic User Equilibrium. *Transportation Science*, 35, 107-123.
https://doi.org/10.1287/trsc.35.2.107.10133

[7] Xu, G., Lam, W. H. K., & Chan, K. S. (2004). Integrated Approach for Trip Matrix Updating and Network Calibration. *ASCE Journal of Transportation Engineering*, 30, 231-244.
https://doi.org/10.1061/(ASCE)0733-947X(2004)130:2(231)

[8] Shafahi, Y. & Faturechi, R. (2009). A New Fuzzy Approach to Estimate the O-D Matrix from Link Volumes. *Transportation Planning and Technology, 32*(6), 499-526.
https://doi.org/10.1080/03081060903374700

[9] Foulds, L. R., Nascimento, H. A. D., Calixto, I. C. A. C., Hall, B. R., & Longo, H. (2013). A Fuzzy Set-based Approach to Origin-Destination Matrix Estimation in Urban Traffic Networks with Imprecise Data. *European Journal of Operational Research*, 231, 190-201.
https://doi.org/10.1016/j.ejor.2013.05.012

[10] Parry, K. & Hazelton, M. L. (2012). Estimation of Origin-Destination Matrices from Link Counts and Sporadic Routing Data. *Transportation Research B*, 46, 175-188.
https://doi.org/10.1016/j.trb.2011.09.009

[11] Ferreira, M. & Damas, L. (2016). Time-Evolving O-D Matrix Estimation Using High-Speed GPS Data Streams. *Expert Systems with Applications*, 44, 275-288.
https://doi.org/10.1016/j.eswa.2015.08.048

[12] Recker, W. W., McNally, M. G., & Root, G. S. (1986). A Model of Complex Travel behaviour: Part I. Theoretical Development. Part II. An Operational Model. *Transportation Research A*, 20, 307-330.
https://doi.org/10.1016/0191-2607(86)90089-0

[13] Pas, E. I. (1983). A Flexible & Integrated Methodology for Analytical Classification of Daily Travel-activity behaviour. *Transportation Science*, 17, 1983, 405-429.
https://doi.org/10.1287/trsc.17.4.405

[14] Huang H. J. & Lam, W. H. K. (2005). A stochastic model for combined activity/destination/route choice problems. *Annals of Operation Research*, 135, 111-125.
https://doi.org/10.1007/s10479-005-6237-5

[15] Axhausen, K. W. (1990). A Simultaneous Simulation of Activity Chains and Traffic Flow. *In: Jones, P., Ed. Development in Dynamic and Activity-Based Approaches to Travel Analysis*, Avebury, Aldershot, England, 206-225.

[16] Fellendorf, M., Haupt, T., Heidl, U., & Scherr, W. (1995). PTV VISION: Activity Based Demand Forecasting in Daily Practices. *The Travel Model Improvement Program Conference*, Daytona Beach, Florida.

[17] Kitamura, R., Pas, E. I., Lula, C. V., Lawton, T. K., & Benson, P. E. (1996). The Sequenced Activity Mobility Simulator SAMS: An Integrated Approach to Modeling Transportation Land Use and Air Quality. *Transportation,* 23, 267-291. https://doi.org/10.1007/BF00165705

[18] Joh, C. H., Timmermans, H. J. P., & Arentze, T. A. (2006). Modelling and Predicting Adaptation Behavior in Multi-dimensional Activity-travel Patterns. *Transportmetrica*, 2, 153-173. https://doi.org/10.1080/18128600608685659

[19] Ramadurai, G. & Ukkusuri, S. (2010). Dynamic user equilibrium model for combined activity-travel choices using activity-travel supernetwork representation. *Networks and Spatial Economics, 10*(2), 273–292.
https://doi.org/10.1007/s11067-008-9078-3

[20] Liao, F., Arentze T., & Timmermans, H. (2013). Incorporating space-time constraints and activity-travel time profiles in a multi-state supernetwork approach to individual activity-travel scheduling. *Transportation Research B*, 55, 41-58. https://doi.org/10.1016/j.trb.2013.05.002

[21] Ouyang, L. Q., Lam, W. H. K., Li, Z. C., & Huang, D. (2011). Network user equilibrium model for scheduling daily activity travel patterns in congested networks. *Journal of the Transportation Research Board*, 2254, 131-139.
https://doi.org/10.3141/2254-14

[22] Liu, P., Liao, F., Huang, H. J., & Timmermans, H. (2015). Dynamic activity-travel assignment in multi-state supernetworks. *Transportation Research B*, 7, 24-42.
https://doi.org/10.1016/j.trb.2015.08.006

[23] Facchinei, F., & Pang, J. S. (2003). *Finite-dimensional Variational Inequalities and Complementarity Problems*. Springer, New York.

[24] Stančić, A., Grgurević, I., & Kavran, Z. (2018). Steganographic Integration of the Data Collected on the Traffic Road. *Tehnički vjesnik, 25*(1), 245-254.
https://doi.org/10.17559/TV-20151228192623

[25] Cascetta, E. (1984). Estimation of Trip Matrices from Traffic Counts and Survey Data: A Generalized Least Squares Estimator. *Transportation Research B,* 18, 289-299.
https://doi.org/10.1016/0191-2615(84)90012-2

[26] Erenoglu, R. C. 2018). A Novel Robust Scaling for EDM Calibration Baselines using Monte Carlo Study. *Tehnički vjesnik, 25*(1), 92-99.
https://doi.org/10.17559/TV-20160407214150

[27] Chootinan, P., Chen, A., & Yang, H. (2005). A Bi-objective Traffic Counting Location Problem for Origin-destination Trip-table Estimation. *Transportmetrica*, 1, 65-80. https://doi.org/10.1080/18128600508685639

Contact information:

Lian Qun OUYANG, Postdoctoral fellow
School of Traffic and Transportation, Beijing Jiaotong University,
No. 3 Shangyuancun, Haidian District, Beijing 100044, P. R. China
E-mail: lillian_auyueng@163.com

Di HUANG, Associate Professor
Corresponding author
Beijing Key Laboratory of Logistics Management and Technology,
Beijing Jiaotong University,
No. 3 Shangyuancun, Haidian District, Beijing 100044, P. R. China
E-mail: huangd@bjtu.edu.cn