



A select link analysis method based on logit–weibit hybrid model

Pengjie Liu¹ · Xiangdong Xu¹ · Anthony Chen^{1,2} · Chao Yang¹ · Longwen Xiao³

Received: 6 December 2016/Revised: 6 July 2017/Accepted: 13 July 2017/Published online: 4 September 2017
© The Author(s) 2017. This article is an open access publication

Abstract Select link analysis provides information of where traffic comes from and goes to at selected links. This disaggregate information has wide applications in practice. The state-of-the-art planning software packages often adopt the user equilibrium (UE) model for select link analysis. However, empirical studies have repeatedly revealed that the stochastic user equilibrium model more accurately predicts observed mean and variance of choices than the UE model. This paper proposes an alternative select link analysis method by making use of the recently developed logit–weibit hybrid model, to alleviate the drawbacks of both logit and weibit models while keeping a closed-form route choice probability expression. To enhance the applicability in large-scale networks, Bell’s stochastic loading method originally developed for logit model is adapted to the hybrid model. The features of the proposed method are twofold: (1) unique O–D-specific link flow pattern and more plausible behavioral realism attributed to the hybrid route choice model and (2) applicability in large-scale networks due to the link-based stochastic loading method. An illustrative network example and a case study in a large-scale network are conducted to demonstrate the efficiency and effectiveness of the proposed select link analysis method as well as applications of O–D-specific link flow information. A visualization

method is also proposed to enhance the understanding of O–D-specific link flow originally in the form of a matrix.

Keywords Select link analysis · Logit model · Weibit model · Hybrid model · Bell loading

1 Introduction

Select link analysis provides information of where traffic comes from and goes to at selected links, i.e., the spatial distribution and origin–destination (O–D) pair composition of aggregate link flow. This disaggregate information has wide applications in practice. For example, this disaggregate information enables to quantify or predict the impact of transportation planning, operations, control and management schemes from multiple perspectives such as congestion alleviation, environmental sustainability and inequity. In detail, we need this information to examine the spatial impacts of a proposed road improvement project. Although the License Plate Recognition system can track partial paths of vehicles to analyze link flow composition, it is still unable to cover the whole transportation network. More importantly, the traffic condition after implementing a transportation planning or management scheme cannot be observed beforehand. Moreover, the state-of-the-art planning software packages often adopt the user equilibrium (UE) model for select link analysis. It is well known that the O–D-specific link flow in the UE model is non-unique, hence requiring an additional criterion for selecting a reasonable solution. To resolve this dilemma, Lu and Nie [1] explored continuity of the UE link flow, proposed a general resolution based on the maximum entropy UE and decomposed the total flow of a link by O–D pairs to examine the spatial impacts of a transportation project.

✉ Xiangdong Xu
xiangdongxu@tongji.edu.cn

¹ The Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai, China

² Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong

³ School of Traffic and Transportation Engineering, Central South University, Changsha, China

Later on, Bar-Gera et al. [2] proposed an additional condition of proportionality for select link analysis. On the other hand, empirical studies have repeatedly revealed that the stochastic user equilibrium (SUE) model more accurately predicts observed mean and variance of choices than the UE model (e.g., [3]). Hence, we aim to use a more accurate probabilistic route choice model to develop a new select link analysis method with both behavioral realism and large-scale network applicability.

In the literature, there are two main types of closed-form probabilistic route choice models: logit [4] and weibit [5]. As to the logit model, Dial [4] developed a link-based stochastic loading method with a forward pass and a backward pass. Van Vliet [6] derived the link choice probability expression according to link weights based on Dial's algorithm. Bell [7] proposed two logit assignment methods as alternatives to Dial's algorithm without path enumeration. Akamatsu [8] utilized the decomposition of path choice entropy to transform path-based into link-based logit assignment. Owing to the homogeneity of variance in the logit model, Nakayama [9] and Nakayama and Chikaraishi [10] presented a q-generalization method based on generalized extreme-value distribution to the random component of utility to allow heteroscedastic variance. Ahipaşaoğlu et al. [11] also introduced a marginal distribution model to relax independently and identically distributed random variance in the logit model and applied the method of successive average to compute flexible traffic flows. As to the weibit model, Kitthamkeson and Chen [12] derived an analytical closed-form expected perceived travel cost to replace the mathematical programming formulation in the weibit SUE model. Sharifi et al. [13] further proposed a modified Dial's STOCH algorithm for the solution. Although with closed-form expressions and wide applications, both logit and weibit models have some drawbacks. First of all, the logit model cannot distinguish the short and long networks with the same absolute route cost difference, and the weibit model cannot distinguish the short and long networks with the same relative route cost difference. In general, routes in a short network (one containing short O–D pair lengths) have lower costs than routes in a long network (one containing long O–D pair lengths). Besides, Yao and Chen [14] made a detailed paradox analysis about the logit, weibit and hybrid models (the hybrid model is just a simple weighted-sum model combining the logit and weibit models) and found that the paradox area of the hybrid model is a combination of the paradox areas of the logit and weibit models. Recently, Xu et al. [15] developed a logit–weibit hybrid model with a closed-form route choice probability expression and equivalent SUE mathematical program formulation to alleviate the above drawbacks of both logit and weibit models simultaneously.

Specifically, this paper makes use of the advanced logit–weibit hybrid model as the underlying route choice behavior model to develop a new select link analysis method with both behavioral realism and computational tractability. As to the *behavioral realism*, (1) the logit–weibit hybrid model has a unique select link analysis result (i.e., O–D pair-specific link flow, or the link flow composition), which avoids the non-uniqueness issue and challenge existing in the current user equilibrium (UE)-based select link analysis; (2) as mentioned earlier, empirical studies (e.g., [3]) have repeatedly revealed that the SUE model more accurately predicts observed mean and variance of choices than the UE model; (3) the logit–weibit hybrid model has behavioral advantages compared to the widely used logit and weibit models. *Computational tractability* is achieved by developing a link-based Bell-type loading algorithm for the hybrid model to guarantee the applicability in large-scale networks. Along with a strict mathematical proof, we modify the calculation of link weights in Bell's [7] link-based algorithm for the logit model to solve the hybrid model, while inheriting its advantages such as easy implementation and avoidance of the high computational expense of path enumeration/storing and the inconsistency issue of path generation.

The remainder of this paper is organized as follows: Sect. 2 briefly introduces the logit–weibit hybrid model and provides a numerical example to compare logit, weibit, and hybrid models. To obtain more detailed link flow information for select link analysis, Sect. 3 presents a link-based solution algorithm for the hybrid model along with the theoretical proof and explores an extension of route overlapping. Two numerical examples are provided in Sect. 4 to demonstrate the algorithm and its by-product (e.g., the choice probability or flow on a segment constituted by several consecutive links by propagating consecutive link conditional choice probabilities). Finally, the application and visualization in a large-scale network are shown in Sect. 5.

2 The logit–weibit hybrid route choice model

The core of select link analysis is the travelers' route choice behavior model. Next, we give a brief introduction to the logit, weibit and hybrid models.

The logit route choice probability expression is as follows [4]:

$$p_k^{rs} = \frac{\exp(-\theta c_k^{rs})}{\sum_{p \in K^{rs}} \exp(-\theta c_p^{rs})} = \frac{1}{\sum_{p \in K^{rs}} \exp[-\theta(c_p^{rs} - c_k^{rs})]}, \forall k \in K^{rs}, h_{rs} \in H, \quad (1)$$

where p_k^{rs} is the choice probability of route k between O–D pair (r, s) ; h_{rs} represents the O–D pair (r, s) ; H is the set of O–D pairs; K^{rs} is the set of routes between O–D pair (r, s) ; c_k^{rs} and c_p^{rs} are the travel time of routes k and p , respectively, between O–D pair (r, s) ; θ is the dispersion parameter.

According to Castillo et al. [5], the weibit route choice probability expression is as follows:

$$p_k^{rs} = \frac{(g_k^{rs})^{-\beta}}{\sum_{p \in K^{rs}} (g_p^{rs})^{-\beta}} = \frac{1}{\sum_{p \in K^{rs}} (g_p^{rs}/g_k^{rs})^{-\beta}}, \quad (2)$$

$\forall k \in K^{rs}, h_{rs} \in H,$

where g_k^{rs} and g_p^{rs} are the travel cost of routes k and p , respectively, between O–D pair (r, s) ; and β is the shape parameter.

The logit–weibit hybrid model proposed by Xu et al. [15] can effectively identify the short and long networks with the identical absolute cost difference as well as with the identical relative cost difference. Its route choice probability expression is as follows:

$$p_k^{rs} = \frac{\exp(-\theta c_k^{rs}) \cdot (g_k^{rs})^{-\beta}}{\sum_{p \in K^{rs}} \exp(-\theta c_p^{rs}) \cdot (g_p^{rs})^{-\beta}}$$

$$= \frac{1}{\sum_{p \in K^{rs}} \exp[-\theta(c_p^{rs} - c_k^{rs})] \cdot (g_p^{rs}/g_k^{rs})^{-\beta}}, \quad (3)$$

$\forall k \in K^{rs}, h_{rs} \in H.$

The hybrid model can be considered a bi-criteria route choice problem, i.e., travel time c and travel cost g . Travel time has an additive structure; i.e., route travel time is the sum of link travel time on that route. Travel cost has a multiplicative cost structure; i.e., route travel cost is a multiplication of link travel cost on that route. Hensher et al. [16] presented an exponential function form relating link cost and link time:

$$\tau_a = e^{0.075c_a}, \quad \forall a \in A, \quad (4)$$

where τ_a and c_a are the travel cost and travel time of link a , respectively. From the denominator of Eq. (3), one can see that the hybrid model considers the absolute route cost difference in the exponential function and the relative route cost difference in the power function.

Next, a simple network (see Fig. 1) consisting of two parallel paths/links is given to compare the logit, weibit, and logit–weibit hybrid models. Assume that $\theta = 0.1$, and $\beta = 2.1$. We consider four cases, which correspond to the absolute cost difference and relative cost difference in the long network, and the absolute cost difference and relative cost difference in the short network, respectively. In case 1 and case 2, the short and long networks have the identical absolute cost difference of 10 units. In the short network, the upper route cost is double that of the lower route cost, while it is only 9% greater in the long network. In case 3 and case 4, the short and long networks have the same identical relative cost difference (i.e., 2 times). In the short network, the upper route cost is 10 units larger than the lower route cost, while it is 100 units larger in the long network. Table 1 gives the probabilities of choosing the lower route from the above three models.

In cases 1 and 2, the logit model gives the same route choice probability for both short and long networks. On the contrary, in cases 3 and 4, the weibit model gives the same results for both short and long networks. Hence, the logit model cannot distinguish the short and long networks with the same absolute route cost difference, and the weibit model is unable to identify the short and long networks with the same relative route cost difference. However, in

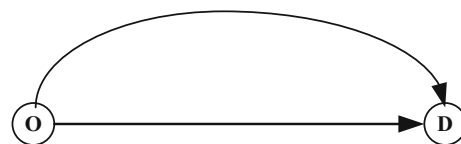


Fig. 1 A simple network consisting of two parallel paths/links

Table 1 Comparison of route choice probability of the logit, weibit, and hybrid models

Case no.	Cost	Probability of choosing the lower route		
		Logit	Weibit	Logit–weibit
1	Upper: 20 Lower: 10	$\frac{e^{-0.1 \times 10}}{e^{-0.1 \times 10} + e^{-0.1 \times 20}} = 0.73$	$\frac{10^{-2.1}}{10^{-2.1} + 20^{-2.1}} = 0.81$	$\frac{e^{-0.1 \times 10} \times 10^{-2.1}}{e^{-0.1 \times 10} \times 10^{-2.1} + e^{-0.1 \times 10} \times 10^{-2.1}} = 0.92$
2	Upper: 110 Lower: 100	$\frac{e^{-0.1 \times 100}}{e^{-0.1 \times 100} + e^{-0.1 \times 110}} = 0.73$	$\frac{100^{-2.1}}{100^{-2.1} + 110^{-2.1}} = 0.55$	$\frac{e^{-0.1 \times 100} \times 100^{-2.1}}{e^{-0.1 \times 100} \times 100^{-2.1} + e^{-0.1 \times 110} \times 110^{-2.1}} = 0.77$
3	Upper: 20 Lower: 10	$\frac{e^{-0.1 \times 10}}{e^{-0.1 \times 10} + e^{-0.1 \times 20}} = 0.73$	$\frac{10^{-2.1}}{10^{-2.1} + 20^{-2.1}} = 0.81$	$\frac{e^{-0.1 \times 10} \times 10^{-2.1}}{e^{-0.1 \times 10} \times 10^{-2.1} + e^{-0.1 \times 10} \times 10^{-2.1}} = 0.92$
4	Upper: 200 Lower: 100	$\frac{e^{-0.1 \times 100}}{e^{-0.1 \times 100} + e^{-0.1 \times 200}} = 0.99$	$\frac{100^{-2.1}}{100^{-2.1} + 200^{-2.1}} = 0.81$	$\frac{e^{-0.1 \times 100} \times 100^{-2.1}}{e^{-0.1 \times 100} \times 100^{-2.1} + e^{-0.1 \times 200} \times 200^{-2.1}} = 1.00$

cases 1, 2 and 4, the hybrid model gives different route choice probabilities, which is more receivable than the other two models. Travelers perceive the 10 units of absolute cost difference to be more significant in a short network (e.g., 10 and 20 min) than in a long network (e.g., 100 and 110 min). Similarly, travelers perceive the 2 times of relative cost difference to be more significant in a long network (e.g., 100 and 200 min) than in a short network (e.g., 10 and 20 min). Hence, the hybrid model can distinguish the short and long networks with the identical absolute cost difference as well as the identical relative cost difference. Namely, the hybrid model can alleviate the drawbacks of the logit and weibit models and is more flexible.

In summary, the hybrid model can effectively identify the absolute difference and relative cost difference of the long or short network, and it has a larger applicability and flexibility than the other two models. It should be noted that the above analysis supposes the logit and hybrid models have the same θ value, and the weibit and hybrid models have the same β value. This setting is only for simplicity, which is not an assumption or requirement of the hybrid model. The main purpose of the above example is to analyze the drawbacks of the logit and weibit models and the capability of the hybrid model in alleviating the drawbacks.

3 Link-based solution algorithm

Note that if we use Eq. (3) to solve the hybrid model directly, we need to enumerate routes or generate a route set, which is a non-trivial task for large-scale networks. Hence, in order to embed this hybrid route choice model into select link analysis, we need to develop an effective solution algorithm with large-scale network applicability for the logit–weibit hybrid model. Specifically, we modify Bell's link-based algorithm (originally proposed for the logit model) for solving the hybrid model. The main modification is on the calculation of link weights. With this, the modified link-based solution algorithm is able to generate the hybrid model-based flow pattern while inheriting the advantages of Bell's algorithm, such as easy implementation and avoidance of path enumeration/storing.

3.1 Algorithm procedure

Step 1: Weights matrix initialization The node-pair weight is initialized as follows:

$$w_{ij} = \begin{cases} \exp(-\theta c_{ij}) \tau_{ij}^{-\beta}, & \forall l_{ij} \in L \\ 0, & \text{otherwise} \end{cases}, \quad (5)$$

where l_{ij} denotes the link ij (i.e., the link from node i to node j), and L is the set of links in the network; w_{ij} denotes

the initial node-pair weight of link l_{ij} ; c_{ij} and τ_{ij} are the travel time and travel cost (see Eq. (4)) of link l_{ij} , respectively. For each link, the corresponding node-pair (from tail node to head node) weight is a function of link travel time and link travel cost. For disconnected node pairs, the weight is 0. Note that herein we use a compound function as the initialized link weight, where the exponential and power functions correspond to the logit and weibit models, respectively.

Step 2: Weights matrix update There are two ways to update the node-pair weight matrix, which are the same as Bell's algorithm [7].

Alternative I: for all nodes m

for all nodes i not equal to m

for all nodes j not equal to m or i

$$w_{ij} = w_{ij} + w_{im} w_{mj}. \quad (6)$$

Alternative II: for initialized weights matrix:

$$\mathbf{W} = (\mathbf{I} - \mathbf{W})^{-1} - \mathbf{I}, \quad (7)$$

where \mathbf{I} is an identity matrix with the same dimension as weights matrix. After processing weights matrix by any of the above two methods, w_{ij} is assigned to be 1.

Step 3: Link choice probability The probability that a trip from node r to node s chooses link l_{ij} can be calculated as follows:

$$p_{ij}^{rs} = \frac{w_{ri} \exp(-\theta c_{ij}) \tau_{ij}^{-\beta} w_{js}}{w_{rs}}, \quad (8)$$

where p_{ij}^{rs} is the probability of choosing link l_{ij} from node r to node s .

Step 4: Link flow The O–D pair-specific link flow is then calculated as

$$v_{ij}^{rs} = p_{ij}^{rs} q_{rs}, \quad (9)$$

where q_{rs} is travel demand of O–D pair (r, s) , and v_{ij}^{rs} is flow on link l_{ij} assigned by O–D pair (r, s) . For a network with multiple O–D pairs, we can get the total link flows by accumulating the link flow assigned by each O–D pair.

Step 5: Select link analysis We can get the total flow of each link as well as its O–D pair composition (i.e., the link flow composition from each O–D pair). In other words, we can get information of where flow comes from and goes to at a selected link, i.e., the spatial distribution and O–D pair composition of aggregate link flow.

3.2 Algorithm modification statement

The above algorithm for the hybrid model is modified on the basis of Bell's [7] link-based algorithm for the logit model. The main modification is the calculation of link weights to cater for the hybrid model. Bell's link-based

algorithm inherits the drawbacks of the logit model and does not take absolute cost difference into account. However, the proposed algorithm contains both absolute cost difference and relative cost difference as shown in Eqs. (5) and (8).

The original weights matrix in Bell [7] is as follows:

$$w_{ij} = \begin{cases} \exp(-\theta c_{ij}), & \forall l_{ij} \in L \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

Compared with Eq. (10), the exponential term in Eq. (5) of the modified algorithm can be regarded as Eq. (10) in [7], and the additional power term (i.e., $\tau_{ij}^{-\beta}$) in Eq. (5) allows the consideration of relative cost difference for the weibit model. Theoretically, the modified algorithm can be applied for solving the hybrid model in large-scale networks without path enumeration/storing/generation.

3.3 Theoretical proof

In a network, when travelers reach node i , the probability of choosing the next node j refers to the conditional probability. Assume travelers on node j have two choices, choosing node j or m . It becomes a conditional probability problem $p(j|i)$, given that node i has been chosen [17]. Suppose link l_{ij} is between a particular O–D pair, we have

$$p(j|i) = \frac{p_{ij}}{p_i} = \frac{w_{ri}w_{ij}w_{js}}{w_{rs}} / \frac{w_{ri}w_{is}}{w_{rs}} = \frac{w_{ij}w_{js}}{w_{is}} \quad (11)$$

Suppose route k has a series of intermediate nodes $m_1, m_2, m_3, \dots, m_n$, the probability of choosing route k between O–D pair (r, s) is given by

$$\begin{aligned} p_k^{rs} &= \prod_{ij} p(j|i)^{\delta_{ij,k}} \\ &= p(m_1|r)p(m_2|m_1)p(m_3|m_2) \cdots p(s|m_n) \\ &= \frac{w_{rm_1}w_{m_1s}}{w_{rs}} \cdot \frac{w_{m_1m_2}w_{m_2s}}{w_{m_1s}} \cdot \frac{w_{m_2m_3}w_{m_3s}}{w_{m_2s}} \cdots \frac{w_{m_ns}w_{ss}}{w_{m_{n-1}s}} \\ &= \frac{w_{rm_1} \cdot w_{m_1m_2} \cdot w_{m_2m_3} \cdots w_{m_ns} \cdot w_{ss}}{w_{rs}} \\ &= \frac{\exp(-\theta c_{rm_1}) \cdot \tau_{rm_1}^{-\beta} \cdot \exp(-\theta c_{m_1m_2}) \cdot \tau_{m_1m_2}^{-\beta} \cdots \exp(-\theta c_{m_ns}) \cdot \tau_{m_{n-1}s}^{-\beta}}{w_{rs}} \\ &= \prod_{ij} \exp(-\theta c_{ij}) \cdot \tau_{ij}^{-\beta} / w_{rs} \\ &= \exp\left(-\theta \sum_{ij} c_{ij}\right) \cdot \left(\prod_{ij} \tau_{ij}\right)^{-\beta} / w_{rs} \\ &= \frac{\exp(-\theta c_k^{rs}) \cdot (g_k^{rs})^{-\beta}}{w_{rs}}, \end{aligned} \quad (12)$$

where $\delta_{ij,k}$ is the number of times that link l_{ij} is used in route k .

Due to the route choice probability conservation between an O–D pair,

$$\sum_{k \in K^{rs}} p_k^{rs} = 1. \quad (13)$$

Substituting Eq. (12) into Eq. (13), we have

$$\sum_{k \in K^{rs}} \frac{\exp(-\theta c_k^{rs}) \cdot (g_k^{rs})^{-\beta}}{w_{rs}} = 1. \quad (14)$$

Then

$$w_{rs} = \sum_{k \in K^{rs}} \exp(-\theta c_k^{rs}) \cdot (g_k^{rs})^{-\beta}. \quad (15)$$

Further substituting Eq. (15) into Eq. (12), we have the following route choice probability expression:

$$p_k^{rs} = \frac{\exp(-\theta c_k^{rs}) \cdot (g_k^{rs})^{-\beta}}{\sum_{p \in K^{rs}} \exp(-\theta c_p^{rs}) \cdot (g_p^{rs})^{-\beta}}, \quad \forall k \in K^{rs}, h_{rs} \in H. \quad (16)$$

In summary, the above improved link-based loading algorithm is able to reach the hybrid route choice model in theory. Hence, it can be used to solve the hybrid route choice model while circumventing the need of path enumeration or generation, which is an important guarantee for large-scale network applications.

3.4 Extension to route overlapping

Even though the above link-based loading algorithm can solve the hybrid model, it still inherits the route independence assumption in the hybrid model. The consequence is that it cannot handle the route overlapping problem. Similar to Russo and Vitetta [18], we can adopt a link-based commonality factor to deal with the route overlapping issue. The commonality factor of link l_{ij} between O–D pair (r, s) can be expressed as follows:

$$CF_{ij}^{rs} = \alpha z_{ij}^{rs} \log(N_{ij}^{rs}), \quad \forall l_{ij} \in L, h_{rs} \in H, \quad (17)$$

where CF_{ij}^{rs} denotes the commonality factor of link l_{ij} between O–D pair (r, s) ; α is a parameter to be calibrated; z_{ij}^{rs} is the ratio between the cost of link l_{ij} and the shortest route cost between O–D pair (r, s) ; N_{ij}^{rs} is the number of routes using link l_{ij} between O–D pair (r, s) . Without loss of generality, we use $\tau_{ij} = e^{\gamma c_{ij}}$ like Kitthamkesorn and Chen [12, 19]. Then, the node-pair weight becomes

$$w_{ij} = \begin{cases} \exp(-\theta(c_{ij} + CF_{ij}^{rs})) \exp(-\gamma\beta(c_{ij} + CF_{ij}^{rs})), & \forall l_{ij} \in L, h_{rs} \in H, \\ 0, & \text{otherwise,} \end{cases} \quad (18)$$

which can also be reorganized as follows:

$$w_{ij} = \begin{cases} \exp(-\theta c_{ij}) \tau_{ij}^{-\beta} \cdot \exp(-(\theta + \gamma\beta)CF_{ij}^{rs}), & \forall l_{ij} \in L, h_{rs} \in H \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

Next, we repeat the above theoretical proof in Sect. 3.3 using Eq. (19). The probability of choosing route k between O–D pair (r, s) is given by

$$\begin{aligned} p_k^{rs} &= \prod_{ij} p(j|i)^{\delta_{jk}} \\ &= p(m_1|r)p(m_2|m_1)p(m_3|m_2) \cdots p(s|m_n) \\ &= \frac{w_{rm_1} w_{m_1s}}{w_{rs}} \cdot \frac{w_{m_1m_2} w_{m_2s}}{w_{m_1s}} \cdot \frac{w_{m_2m_3} w_{m_3s}}{w_{m_2s}} \cdots \frac{w_{m_{n-1}s} w_{ss}}{w_{m_{n-1}s}} \\ &= \prod_{ij} \exp(-\theta c_{ij}) \cdot \tau_{ij}^{-\beta} \cdot \exp(-(\theta + \gamma\beta)CF_{ij}^{rs}) / w_{rs} \\ &= \exp\left(-\theta \sum_{ij} c_{ij}\right) \cdot \left(\prod_{ij} \tau_{ij}\right)^{-\beta} \cdot \exp\left(-(\theta + \gamma\beta) \sum_{ij} CF_{ij}^{rs}\right) / w_{rs}. \end{aligned} \quad (20)$$

Russo and Vitetta [18] also presented the path-based commonality factor as follows:

$$CF_k^{rs} = \alpha \sum_{ij \in k} z_k^{rs} \log\left(N_{ij}^{rs}\right), \quad \forall k \in K^{rs}, h_{rs} \in H, \quad (21)$$

where CF_k^{rs} denotes the commonality factor of route k between pair (r, s) ; z_k^{rs} is the perceived relevance of an individual link l_{ij} in route k . Substituting Eq. (21) into Eq. (20), we have

$$p_k^{rs} = \frac{\exp(-\theta c_k^{rs}) \cdot (g_k^{rs})^{-\beta} \cdot \exp(-(\theta + \gamma\beta)CF_k^{rs})}{w_{rs}}. \quad (22)$$

Referring to Eqs. (13)–(16), we have the following probability expression:

$$\begin{aligned} p_k^{rs} &= \frac{\exp(-\theta c_k^{rs}) \cdot (g_k^{rs})^{-\beta} \exp(-(\theta + \gamma\beta)CF_k^{rs})}{\sum_{p \in K^{rs}} \exp(-\theta c_p^{rs}) \cdot (g_p^{rs})^{-\beta} \exp(-(\theta + \gamma\beta)CF_p^{rs})}, \\ &\forall k \in K^{rs}, h_{rs} \in H. \end{aligned} \quad (23)$$

This is the hybrid route choice probability expression with a route overlapping consideration. It can adjust the probability of routes coupling with other routes by the commonality factor.

4 Numerical examples

This section provides two numerical examples to verify the capability and applicability of the modified link-based stochastic loading algorithm for solving the hybrid model: The first network is a grid network, and the second network is the Sioux Falls network. The parameter values are set as $\theta = 0.35$ and $\beta = 3.7$.

4.1 Grid network

This grid network consists of 9 nodes and 12 links. Each link is represented by its tail node and head node. The link costs are provided in Fig. 2, and there are four O–D pairs, i.e., O–D pairs (1, 9), (2, 9), (4, 9) and (5, 9), with the same traffic demand of 1000 units.

To set up a benchmark, we enumerate all routes of this network, each O–D pair having 6, 3, 3 and 2 routes, respectively, and calculate the route choice probability and flow according to Eq. (3). With these, we can obtain the sources and composition of each link flow (see Table 2). Then, we perform the above link-based loading algorithm and compare the results with the benchmark results obtained by path enumeration. As expected, the link-based loading algorithm can get exactly the same results as the path enumeration approach. This verifies the correctness of the algorithm.

Although Sect. 2 has provided a comparison among the logit, weibit and hybrid models, we take a more detailed illustration of the three models in the grid network with the uniform O–D demands (see Table 3 and Table 4). The value of dispersion parameter θ of the logit and hybrid models is equal, and the value of shape parameter β of the weibit and hybrid models is the same. The grid network can be regarded as a short network. By comparing the assignment results of the three models, we can verify that the hybrid model can distinguish the short networks with the identical absolute cost difference as well as the identical relative cost difference. Figure 3 gives the total link flows assigned by the three models. We can find that the links with larger flows have larger difference among the three models except for link 2–5. Due to the lack of realistic link flow observations, we cannot directly judge which assignment model is more realistic. However, the simultaneous consideration of absolute and relative cost differences makes the hybrid model outperform both the logit and weibit models.

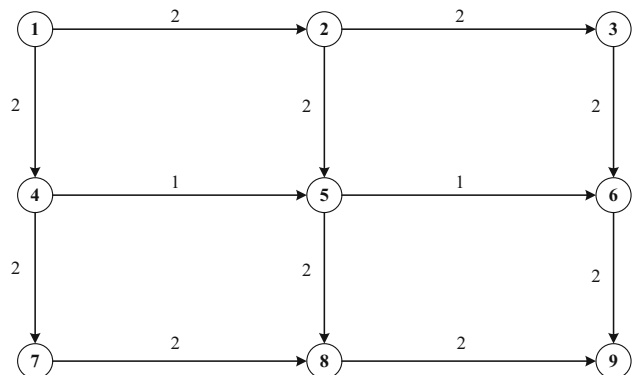


Fig. 2 Grid network with link cost

Table 2 Link flow assigned by O–D pairs and total link flow by the logit–weibit hybrid model

O–D	Links											
	1–2 ^a	1–4	2–3	2–5	3–6	4–5	4–7	5–6	5–8	6–9	7–8	8–9
(1,9)	377.7	622.3	97.5	280.2	97.5	524.8	97.5	524.8	280.2	622.3	97.5	377.7
(2,9)	– ^b	–	258.2	741.8	258.2	–	–	483.6	258.2	741.8	–	258.2
(4,9)	–	–	–	–	–	843.3	156.7	549.8	293.5	549.8	156.7	450.2
(5,9)	–	–	–	–	–	–	–	651.9	348.1	651.9	–	348.1
Total	377.7	622.3	355.7	1022	355.7	1368	254.2	2210	1180	2565.8	254.2	1434.2

^a The expression form of link

^b The O–D pair cannot assign flow to the corresponding link because of the network typology

Table 3 Link flow assigned by O–D pairs and total link flow by the logit model

O–D	Links											
	1–2 ^a	1–4	2–3	2–5	3–6	4–5	4–7	5–6	5–8	6–9	7–8	8–9
(1,9)	435.4	564.6	127.4	308.1	127.4	437.2	127.4	437.2	308.1	564.6	127.4	435.4
(2,9)	– ^b	–	292.5	707.5	292.5	–	–	415.0	292.5	707.5	–	292.5
(4,9)	–	–	–	–	–	774.4	225.6	454.3	320.1	454.3	225.6	545.7
(5,9)	–	–	–	–	–	–	–	586.6	413.4	586.6	–	413.4
Total	435.4	564.6	419.8	1015.6	419.8	1211.6	352.9	1893.1	1334.1	2313	352.9	1687

^a The expression form of link

^b The O–D pair cannot assign flow to the corresponding link because of the network typology

Table 4 Link flow assigned by O–D pairs and total link flow by the weibit model

O–D	Links											
	1–2 ^a	1–4	2–3	2–5	3–6	4–5	4–7	5–6	5–8	6–9	7–8	8–9
(1,9)	449.7	550.3	135.5	314.3	135.5	414.8	135.5	414.8	314.3	550.3	135.5	449.7
(2,9)	– ^b	–	301.2	698.8	301.2	–	–	397.6	301.2	698.8	–	301.2
(4,9)	–	–	–	–	–	753.8	246.2	428.9	324.9	428.9	246.2	571.1
(5,9)	–	–	–	–	–	–	–	568.9	431.1	568.9	–	431.1
Total	449.7	550.3	436.7	1013.1	436.7	1168.6	381.7	1810.1	1371.5	2246.8	381.7	1753.2

^a The expression form of link

^b The O–D pair cannot assign flow to the corresponding link because of the network typology

Below we analyze the traffic sources and composition of each aggregate link flow. Without loss of generality, we consider four O–D pairs: (1, 9), (2, 9), (4, 9) and (5, 9). Figure 4 shows the flow sources of each link from the four O–D pairs. For example, all flows of link 1–2 and link 1–4 come from the O–D pair (1, 9); 2210 flow units on link 5–6 consist of 524.8 units from O–D pair (1, 9), 483.6 units from (2, 9), 549.8 units from (4, 9), and 651.9 units from (5, 9). In practice, this disaggregate information enables to quantify or predict the impact of transportation planning, operations, control and management schemes from

multiple perspectives such as congestion alleviation, environmental sustainability, and equity issue.

As a by-product of the above information, the choice probability of a corridor or a route constituted by several consecutive links can be obtained by propagating consecutive link choice probabilities without path enumeration, which avoids calculating the overall link flows in the context of requiring merely some selected link flows. Note that the above choice probability refers to the conditional probability that travelers choose the next link under the condition of choosing the current link, i.e.,

$$p(l_{jk}|l_{ij}) = \frac{p_{jk}}{\sum_h p_{jh}}, \tag{24}$$

where link l_{jk} is the next link directly connected with link l_{ij} ; $p(l_{jk}|l_{ij})$ is the conditional probability choosing link l_{jk} under the condition of choosing the link l_{ij} ; p_{jk} is the choice probability of link l_{jk} ; p_{jh} is the choice probability of link l_{jh} . After processing link choice probability of each O–D pair, we can get the corridor or route choice probability. For example, the choice probability of corridor 5–6–9 between O–D pair (1, 9) is

$$\begin{aligned} p_{5-6-9} &= p_{5-6}/(p_{5-6} + p_{5-8}) \cdot p_{6-9}/p_{6-9} \\ &= (37.77 \times 28.02 \div (28.02 + 9.75) \\ &\quad + 62.23 \times 52.48 \div (52.48 + 9.75)) \\ &\quad \times 52.48 \div (52.48 + 28.01) \times 1(\%) = 52.48\%. \end{aligned}$$

By multiplying travel demand of the O–D pair by corridor choice probability, we can obtain corridor flow under each O–D pair and consequently the total corridor flow by accumulating all O–D pairs.

4.2 Sioux Falls network

In this section, we apply the link-based loading algorithm to the Sioux Falls network and compare the hybrid model result with those of the logit and weibit models by setting the same parameters as before. Figure 5 shows that the three models appear similar in the Sioux Falls network, and the three curves overlap highly in the view of overall trend. However, there are indeed some small differences, which are expressed more clearly in Figs. 6 and 7. Figure 6 presents the absolute link flow difference, which fluctuates up and down around zero value. For the difference between weibit and hybrid models, the whole curve changes intensely, showing a big difference in a detailed comparison

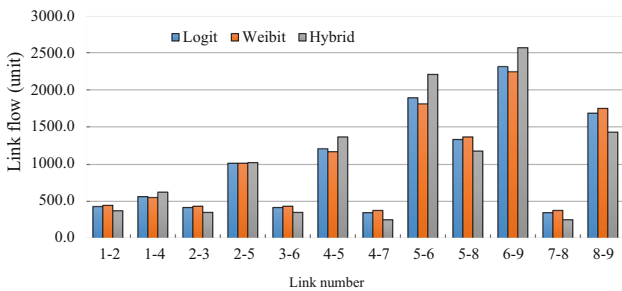


Fig. 3 Total link flows assigned by the logit, weibit and hybrid models

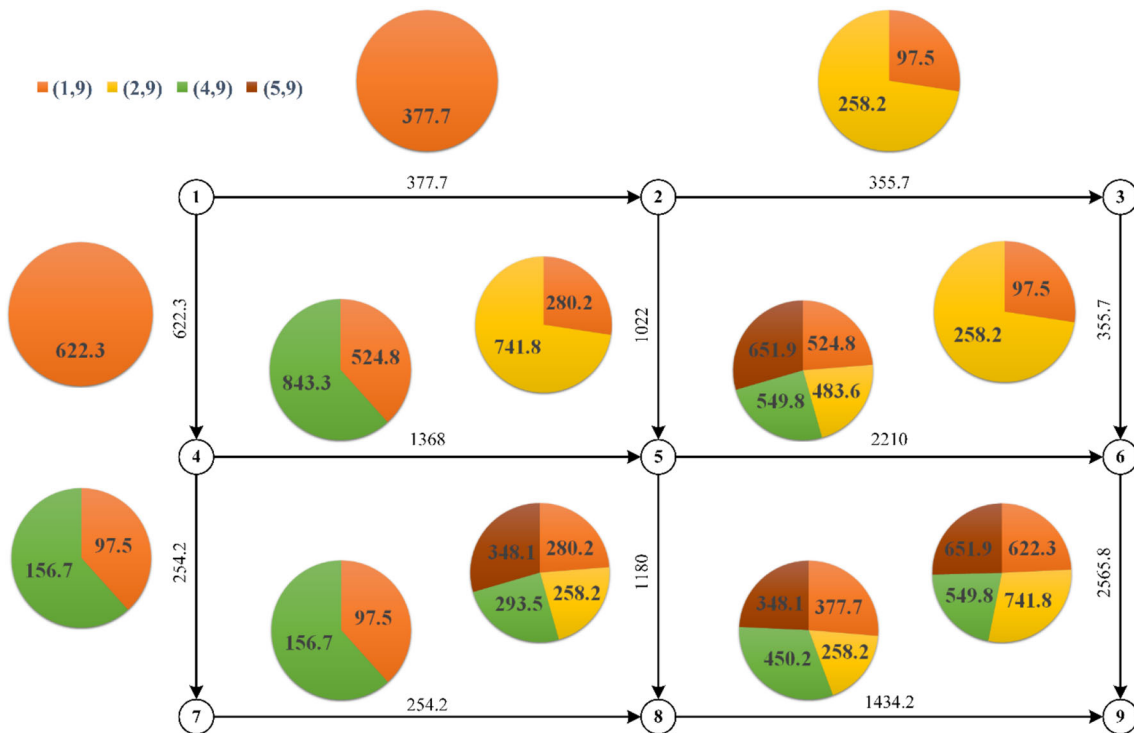


Fig. 4 Sources and composition of each link flow

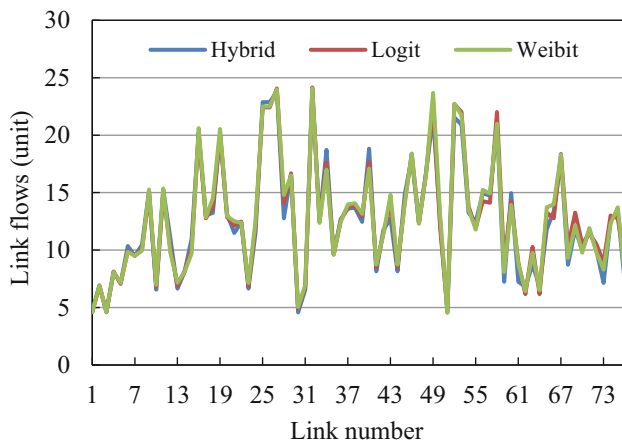


Fig. 5 Link flows assigned by the logit, weibit and hybrid models

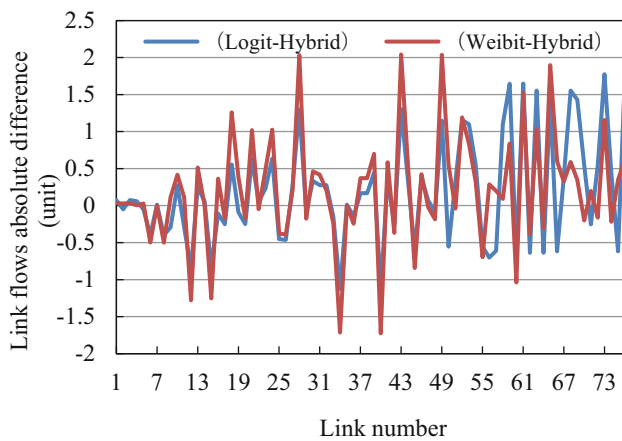


Fig. 6 Absolute difference of link flows between logit/weibit and hybrid models

son. As for the difference between logit and hybrid models, the left chart presents a less change trend than the right part. The relative link flow difference is presented in Fig. 7, which is distributed uniformly around the ratio value of one. The ratios of logit to hybrid models and weibit to hybrid models keep approximately identical except for the right corner part. Combining Figs. 6 with 7, we can find that the traffic assigned by logit and hybrid models is closer than that by weibit and hybrid models.

Next, to obtain specific link flow sources and composition, we analyze link flow composition by taking link 17–19 as an example. Since there are 550 O–D pairs, we only show the O–D pairs whose assigned link flow accounts for not less than 1% of the total link flow. Figures 8 and 9 contain 25 O–D pairs, which occupy at least 1% proportion in flow sources of link 17–19. Among them, four O–D pairs have the largest proportion, i.e., (10, 19), (17, 19), (17, 20) and (17, 22). Some O–D pairs with a smaller proportion are those having node 19 as the

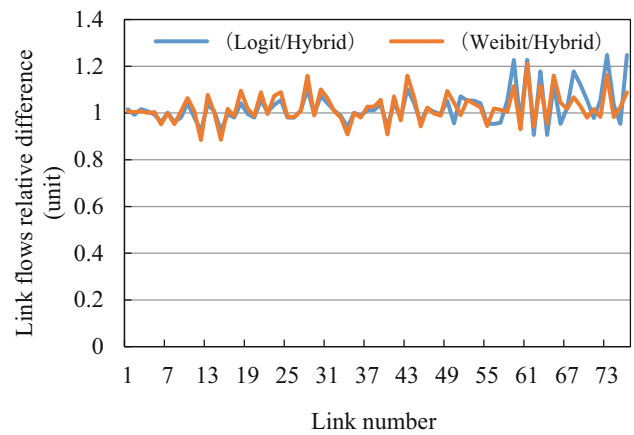


Fig. 7 Relative difference of link flows between logit/weibit and hybrid models

destination. A possible reason is that these O–D pairs do not have large traffic demands. Furthermore, some O–D pairs without direct link connection also occupy a large proportion, such as (7, 15) and (16, 14), which is not apparent without conducting select link analysis.

Below we explore the application of select link analysis in identifying the impact area of a system intervention (e.g., traffic incident and road closure). We close link 17–19 (without constructing alternative roads) to observe the impact on the network and to explore the method of identifying the impact area. There are two possible methods to identify the affected area: the first one is according to the change of total link flows (see Table 5), and the second one is according to the link flow sources before closing link 17–19. Herein, we identify the affected area according to these two methods, respectively.

In Table 5, the positive values represent increase in link flows and negative values represent decrease in link flows after closing link 17–19. We can observe that the majority of the links are affected with different relative change degrees of flows although only link 17–19 is closed; most O–D pairs that are likely to use this link will be affected. In order to identify the affected area, we set the criterion as whether the link flow change is greater than 20%. Namely, those links with more than 20% positive or negative flow change are regarded as affected links. According to this criterion, the impact area of closing link 17–19 is shown by the red circled area in Fig. 10. Combining Table 5 with Fig. 10, we can find that links further away from link 17–19 are almost unaffected, and the major affected area is located surrounding link 17–19. Furthermore, the closer a link is from link 17–19, the greater flow change the link has.

The second identification method is based on the link flow sources and composition before closing link 17–19 as shown in Fig. 9. The impact area includes nodes/zones

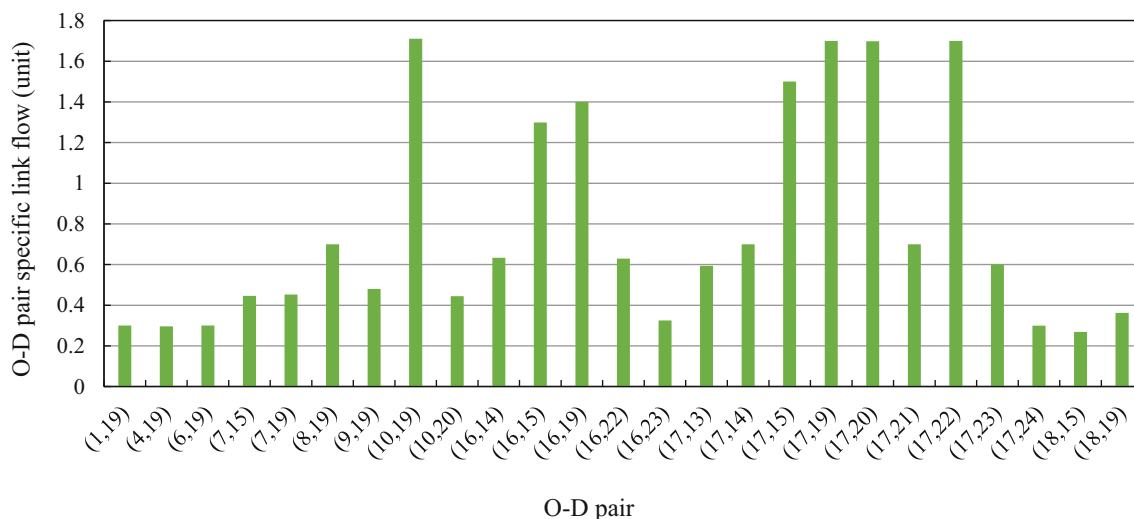


Fig. 8 Link flow sources

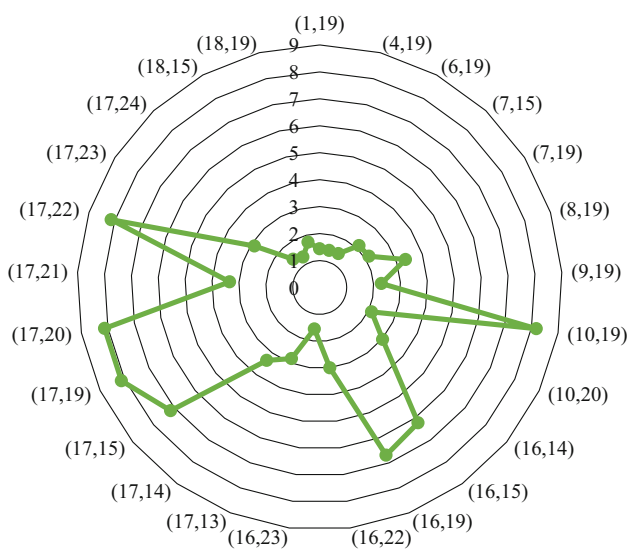


Fig. 9 Link flow sources percentage (%)

whose link flow source percentage is greater than the threshold of 5%. If the percentage is greater than the threshold, the corresponding origin and destination of the link flow source are included in the impact area. The impact area of closing link 17–19 is shown by the circled area in Fig. 11.

Comparing Figs. 10 with 11 reveals that the two impact areas have a little difference, e.g., node 18 is not included into the impact area of the second method (based on the link flow sources and composition before closing link 17–19). By analyzing the identified impact area, we can find that link 17–19 is a central link of the impact area, and links between the set of nodes directly connected to node 17 and node 19 are subject to a greater impact of this link closure. In realistic networks, we can use this analysis to

quantitatively identify the impact area in traffic impact analysis and traffic operations.

5 Application and visualization in a large-scale network

This section demonstrates the applicability and visualization of the proposed link-based loading algorithm and the select link analysis using Chicago sketch network. This network has 386 zones, 933 nodes, 2950 links, and 93,135 O–D pairs. Parameters are set as follows: $\theta = 0.35$ and $\beta = 3.7$. After obtaining each link flow assigned by every O–D pair as well as the total link flows, we can analyze and visualize link flow sources on a selected link.

Figure 12 uses link color and thickness to visualize the assigned total link flows of Chicago sketch network. In Fig. 12, we can get a general and rough cognition on link flows in this network. For example, we can see that traffic is concentrated on the right of the central area, i.e., link flows are clearly larger than the other regions. This is consistent with the actual characteristics of this area, whose road network is denser than other regions. Schools, squares, parks, museums, government agencies and the airport are located in this region, which is a cultural, commercial and municipal land in Chicago. The region generates and attracts a lot of trips, rendering a traffic dense region. In summary, the traffic assignment result is consistent with the road network characteristics.

Chicago sketch network has 2950 links. Among them, link 411–695 is chosen for select link analysis, whose total link flow is 5273 units. Figure 13 presents the proportion of assigned link flow from 21 O–D pairs, whose proportion is greater than 1% of the total link flow. The flow sources of

Table 5 Relative change ratio of link flows after closing link 17–19 (%)

Link	1-2 ^a	1-3	2-1	2-6	3-1	3-4	3-12	4-3	4-5	4-11	5-4
Ratio	-1.30	0.86	0	-0.73	0	-0.22	0.87	0	-2.34	5.08	0
Link	5-6	5-9	6-2	6-5	6-8	7-8	7-18	8-6	8-7	8-9	8-16
Ratio	-5.01	3	0	0	-3.04	0.90	9.41	0	10.37	2.92	-18.42
Link	9-5	9-8	9-10	10-9	10-11	10-15	10-16	10-17	11-4	11-10	11-12
Ratio	0	-0.82	2.75	0	8.64	42.95	-5.46	-29.77	0	-0.64	4.56
Link	11-14	12-3	12-11	12-13	13-12	13-24	14-11	14-15	14-23	15-10	15-14
Ratio	10.81	0	0.46	4.40	0	0.19	-0.01	6.51	-0.46	0	-19.56
Link	15-19	15-22	16-8	16-10	16-17	16-18	17-10	17-16	17-19	18-7	18-16
Ratio	19.95	-20.69	0	20.33	-46.3	78.78	30.36	37.60	^b	0.01	-13.78
Link	18-20	19-15	19-17	19-20	20-18	20-19	20-21	20-22	21-20	21-22	21-24
Ratio	85.53	-57.38	-0.04	-48.72	0	82.22	3.99	35.07	0.16	0.13	-3.85
Link	22-15	22-20	22-21	22-23	23-14		23-24	24-13	24-21	24-23	
Ratio	0.45	0.20	-6.62	0.54	0.39	0.15	-0.71	-4.57	0.17	0.15	

The bold data denotes that the link flow change is greater than 20%

^a The expression form of link

^b The O–D pair cannot assign flow to the corresponding link because of the network topology

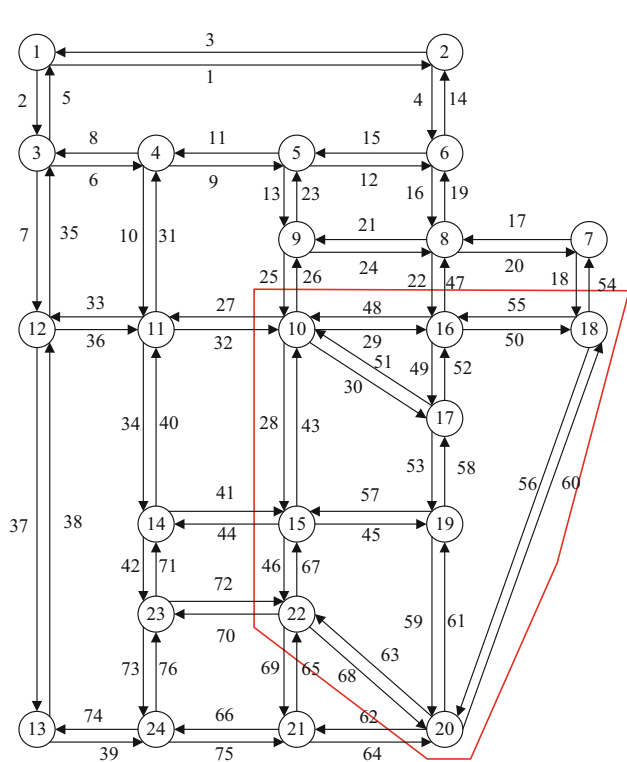


Fig. 10 Impact area identification by using the link flow change as the criterion

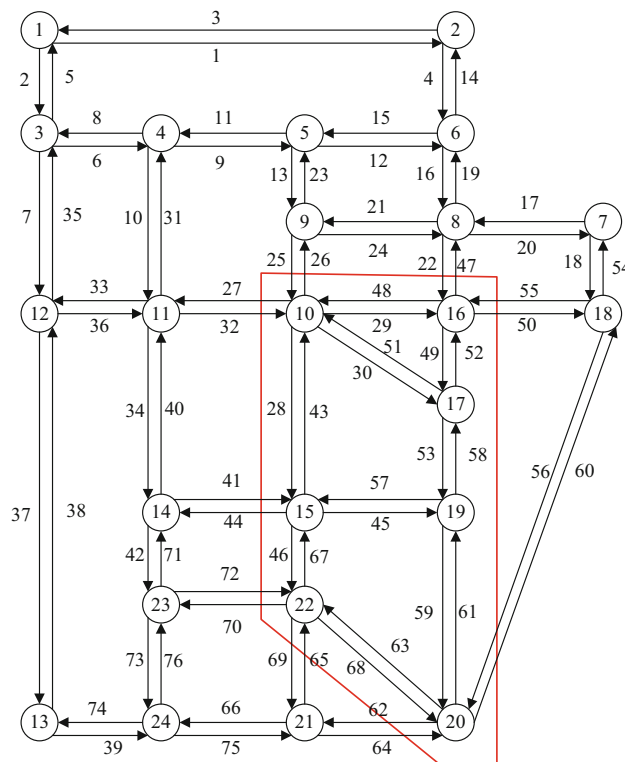


Fig. 11 Impact area identification based on the link flow sources and composition before closing link 17–19

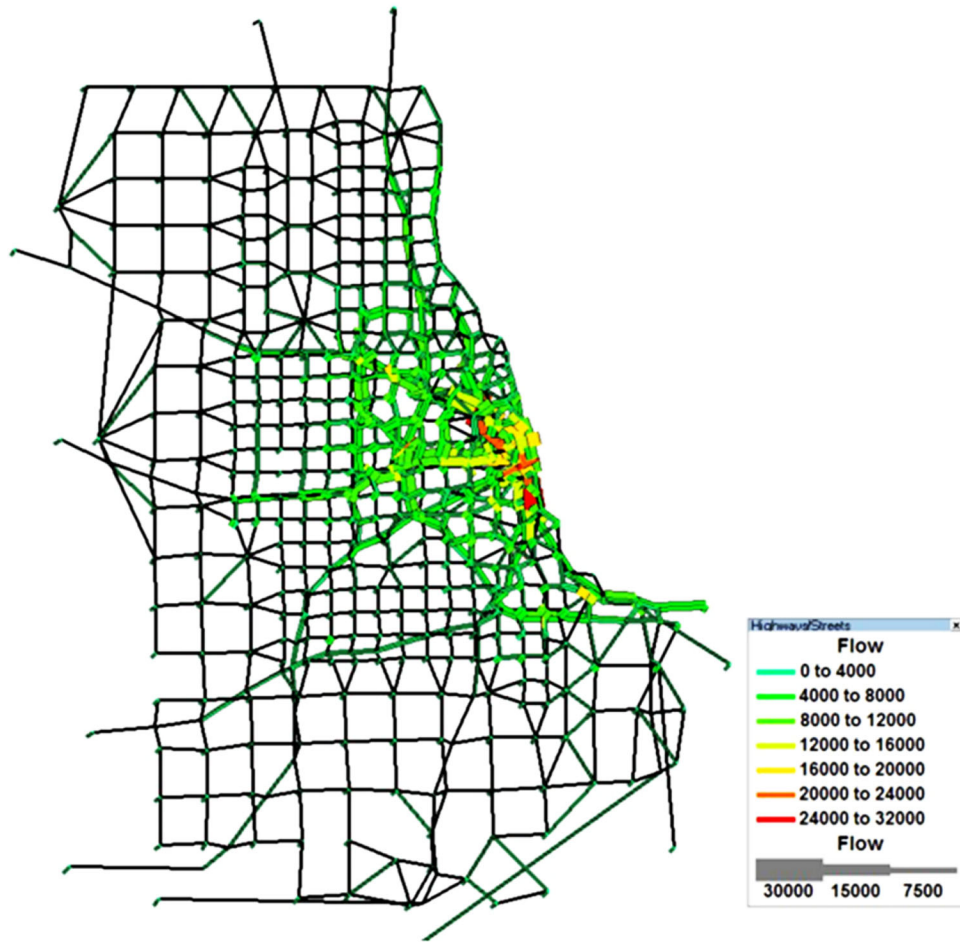


Fig. 12 Total link flows in Chicago sketch network

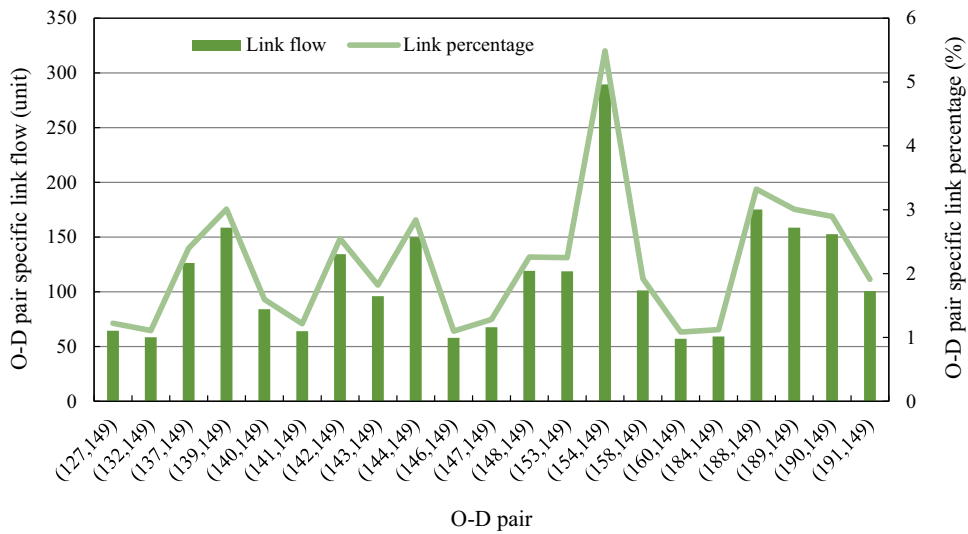


Fig. 13 Link flow sources

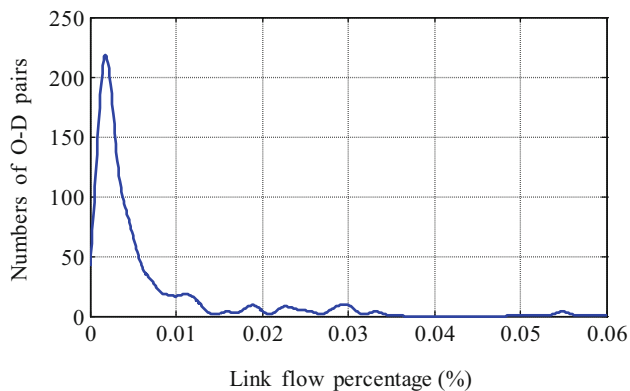


Fig. 14 Probability density distribution of link flow sources

link 411–695 are more concentrated on these 21 O–D pairs, with an accumulative proportion of 45.5% of its total link flow. It has a proportion of about 5.5% in O–D pair (154, 149), and the remaining flows are more evenly distributed in the vicinity of 100 units. Figure 14 shows the entire frequency distribution of link flow sources from all O–D pairs. One can see that the distribution is highly asymmetric with a long distribution tail. The distribution tail represents O–D pair sources with high composition proportions. By combining Fig. 13 with Fig. 14, we can more clearly understand the sources concentration and get specific traffic sources and composition of a selected link.

6 Conclusion

In this paper, we proposed an alternative select link analysis method by using the logit–weibit hybrid model as the underlying probabilistic route choice mechanism. To make it applicable in large-scale networks, Bell’s link-based stochastic loading algorithm originally developed for the classical logit model was modified to solve the hybrid model. The theoretical validity of the modification was rigorously proved. An extension to route overlapping issue was also presented. Three network examples were provided to demonstrate the validity, capability and large-scale network applicability of the link-based algorithm and the select link analysis method. As an application demonstration, we used the information of link flow sources and composition to quantitatively identify the impact area of a road closure. As a future research, we will explore the possible non-convergence issue of the weights matrix in large-scale networks.

Acknowledgements The authors are grateful to the three anonymous referees for their constructive comments and suggestions to improve the quality and clarity of the paper. This study was supported by National Natural Science Foundation of China (51408433), Fundamental Research Funds for the Central Universities of China, the

Chenguang Program sponsored by Shanghai Education Development Foundation and Shanghai Municipal Education Commission.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Lu S, Nie YM (2010) Stability of user-equilibrium route flow solutions for the traffic assignment problem. *Transp Res Part B* 44(4):609–617
- Bar-Gera H, Boyce D, Nie YM (2011) User-equilibrium route flows and the condition of proportionality. *Transp Res Part B* 46(3):440–462
- Dixit VV, Denant-Boemont L (2014) Is equilibrium in transport pure Nash, mixed or Stochastic? *Transp Res Part C* 48:301–310
- Dial R (1971) A probabilistic multipath traffic assignment model which obviates path enumeration. *Transp Res* 5(2):83–111
- Castillo E, Menéndez JM, Jiménez P, Rivas A (2008) Closed form expression for choice probabilities in the Weibull case. *Transp Res Part B* 42(4):373–380
- Van Vliet D (1981) Selected node-pair analysis in Dial’s assignment algorithm. *Transp Res Part B* 15:65–68
- Bell MGH (1995) Alternatives to Dial’s logit assignment algorithm. *Transp Res Part B* 29(4):287–295
- Akamatsu T (1997) Decomposition of path choice entropy in general transport networks. *Trans Sci* 31(4):349–362
- Nakayama S (2013) q-Generalized logit route choice and network equilibrium model. *Procedia-Soc Behav Sci* 80:753–763
- Nakayama S, Chikaraishi M (2015) Unified closed-form expression of logit and weibit and its extension to a transportation network equilibrium assignment. *Transp Res Part B* 81:672–685
- Ahipaşaoğlu SD, Arıkan U, Natarajan K (2016) On the flexibility of using marginal distribution choice models in traffic equilibrium. *Transp Res Part B* 91:130–158
- Kitthamkesorn S, Chen A (2014) Unconstrained weibit stochastic user equilibrium model with extensions. *Transp Res Part B* 59:1–21
- Sharifi MS, Chen A, Kitthamkesorn S, Song Z (2015) Link-based stochastic loading methods for the weibit route choice model. *J Transp Res Rec* 2497:84–94
- Yao J, Chen A (2014) An analysis of logit and weibit route choices in stochastic assignment paradox. *Transp Res Part B* 69:31–49
- Xu X, Chen A, Kitthamkesorn S, Yang H, Lo HK (2015) Modeling absolute and relative cost differences in stochastic user equilibrium problem. *Transp Res Part B* 81(3):686–703
- Hensher DA, Truong TP (1985) Valuation of travel times savings. *J Transp Econ Policy* 19(3):237–260
- Cheng L, Wang W (2002) The research of Dial’s traffic assignment model and choice probability. *J Transp Syst Eng Inf Technol* 2(3):29–32
- Russo F, Vitetta A (2003) An assignment model with modified logit, which obviates enumeration and overlapping problems. *Transportation* 30:177–201
- Kitthamkesorn S, Chen A (2013) A path-size weibit stochastic user equilibrium model. *Transp Res Part B* 57:378–397