Joint optimisation of transfer location and capacity for a capacitated multimodal transport network with elastic demand: a bi-level programming model and paradoxes

Ye, Jiao; Jiang, Yu; Chen, Jun; Liu, Zhiyuan; Guo, Renzhong

Published in:
Transportation Research. Part E: Logistics and Transportation Review

Link to article, DOI:
10.1016/j.tre.2021.102540

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Joint optimisation of transfer location and capacity for a capacitated multimodal transport network with elastic demand: a bi-level programming model and paradoxes

Jiao Ye\textsuperscript{a}, Yu Jiang\textsuperscript{b,*,}\textsuperscript{*}, Jun Chen\textsuperscript{c,d,*}, Zhiyuan Liu\textsuperscript{c,d}, Renzhong Guo\textsuperscript{a}

\textsuperscript{a} Research Institute for Smart Cities Shenzhen University, Shenzhen, Guangdong 518060, China
\textsuperscript{b} DTU Transport, Department of Technology, Management and Economics Technical University of Denmark, 2800, Denmark
\textsuperscript{c} School of Transportation Southeast University, Nanjing, Jiangsu 210096, China
\textsuperscript{d} Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies Southeast University, Nanjing, Jiangsu 210096, China

ARTICLE INFO

Keywords:
Multimodal transport
Bi-level programming
Network design
Paradox

ABSTRACT

With the growing attention toward developing a multimodal transport system to enhance urban mobility, there is an increasing need to construct new infrastructures, rebuild or expand the existing ones, to accommodate the current and newly generated travel demand. Therefore, this study develops a bi-level model that simultaneously determines the location and capacity of the transfer infrastructure to be built considering the elastic demand in a multimodal transport network. The upper-level problem is formulated as a mixed-integer linear programming problem, whereas the lower-level problem is a combined trip distribution/modal split/assignment model that depicts both the destination and route choices of travellers via a multinomial logit model. Numerical studies are conducted to demonstrate the occurrence of two Braess-like paradox phenomena in a multimodal transport network. The first one states that under fixed demand, constructing new parking spaces to provide the usage of park-and-ride services could deteriorate the system performance measured by the total passengers’ travel time, while the second one reveals that under elastic demand, increasing the parking capacity for park-and-ride services to promote its usage may fail, which would be represented by the decline in their modal share. Meanwhile, a numerical example also suggests that constructing transfer infrastructures at distributed stations outperforms building a large transfer centre in terms of attracting travellers using sustainable transit modes.

1. Introduction

Prompting urban multimodal mobility and the use of public transport has attracted increasing attention from both industry and academia, as it has been regarded as an effective strategy to mitigate severe urban congestion (Pineda et al., 2016; Zhao et al., 2019). One of the critical factors measuring the efficiency and attractiveness of a multimodal transport network is whether travellers can seamlessly egress from one subnetwork to access another, which is determined by the integration of different transport modes. For integrated multimodal transport services, increasing attention has been paid to the development of Intelligent Technology (IT) solutions (Xu et al., 2021). From our perspective, the development of IT solutions should also be driven by advances in modelling...
methodologies and built upon the understanding of multimodal networks. Thus, mathematical models still need to be established to facilitate the decision-maker to gain insight into the properties of the multimodal networks to develop effective IT solutions better. Hence, this study is motivated to model and solve one of the fundamental problems for integrating multimodal transport services. That is how to add new transfer links or expand transfer capacities, which is classified as the multimodal network design problem (Farahani et al., 2013). Furthermore, it is well known that a similar problem in a road transport network is alluded to the Braess paradox, which is a counterintuitive phenomenon stating that adding a new link or expanding the capacity of an existing link may not improve the network performance measured by the total travel cost in the network design problem. Notwithstanding, in the context of the multimodal network design problem, the paradoxical phenomenon has not been examined, though expected as a result of the contradiction between the user equilibrium and the system optimum. Thus, another motivation of this study is to demonstrate the characteristics of the paradox phenomenon in a multimodal transport network, as exploring the conditions for the occurrence of Braess paradox could help avoid Braess paradox and improve the system efficiency (Yao et al., 2019a).

With regards to the network design problem, existing studies can be classified into three categories in general based on their decision variables and how to model travellers’ behaviour, that is, mode and route choices. In the first category, the decision variables are associated with each transport mode (i.e., bus route structure, frequency, or road link capacity) and are exclusive to each transport mode (Szeto and Jiang, 2012; Yao et al., 2012; Yan et al., 2013; Tang et al., 2020). This allows executing a traffic/transit assignment model independently on each transport subnetwork and employs a symmetric link performance function, that is, the BPR function (e.g., Beltran et al., 2009; Chen et al., 2020a,b; Jiang and Szeto, 2016; Lee and Vuchic, 2005; Szeto et al., 2011, 2015; Wang et al., 2021b). This category overlooks the fact that the decision variables of one transport mode could affect the network topology of another transport mode, and different transport modes could compete for the same limited infrastructure resources. For example, when buses and cars share the same lane, the construction or expansion of a bus lane leads to capacity reduction for cars, making transit lanes ineffective sometimes. This issue is addressed by the second category, which designs the allocation of infrastructure resources in a multimodal context (e.g., Elshafei, 2006; Mesbah et al., 2008; Li and Ju, 2009; Fan and Machemehl, 2011; Yao et al., 2015; Yu et al. 2015; Huang et al., 2021). The above two categories assume that travellers’ modes and route choices are dedicated to one specific transport mode. In other words, travellers’ trips only contain one travel model1, that is, car or public transport, which overlooks the intermodal travel behaviour, meaning that a traveller could utilise more than one transport mode to complete a trip, such as park-and-ride (P+R) and bike-and-ride (B+R). In the last decade, with the development of multimodal mobility, intermodal travel has become a prevailing option for commuters. In view of this, the third category emerged. In other words, the attributes of transfer nodes in a multimodal transport network (Alumur et al., 2012; Liu and Meng, 2014; Huang et al., 2020; Chen et al., 2016). This category captures a more realistic travel behaviour in which intermodal trips are considered in the route choices. Nevertheless, to the best of our knowledge, studies within this category are limited.

Irrespective of the modelling categories, a common methodology for formulating the multimodal network design problem is bi-level programming. For the upper-level problem, most of the existing studies on multimodal transport network design focus on the location, capacity, or parking fee of the P + R facilities (Liu et al., 2021; Zhong et al., 2020; Chen et al., 2021; Liu et al., 2021), while the B + R facilities are ignored. It has been acknowledged that the B + R facilities could provide access/egress services for commuters who travel between 2 and 5 km to a public transport stop (Martens, 2004). Such transfer infrastructure has the advantages of less area occupancy and less impact on the nearby network. However, there is no existing study examining the simultaneous determination of the location and capacity of multiple types of transfer infrastructure considering the B + R facilities. Furthermore, most studies on P + R services declare that setting P + R facilities would shift commuters from the auto mode to transit and P + R modes (Wang et al., 2004; Liu et al., 2009; Wang et al., 2014; Pineda et al., 2016). Nevertheless, these studies were conducted under a fixed total travel demand; the impacts under variable demand have not been examined before. For the lower-level problem, a combined mode-split and traffic-assignment model is commonly employed. Abdulaal and LeBlanc (1979) proposed an equilibrium condition in which it is assumed that all travel modes are available to commuters when they plan their route. Meng and Liu (2012) adopted a similar condition to formulate a combined mode split and traffic assignment problem to design a multimodal toll pattern in a bimodal network. Xu et al. (2018) considered the combined mode and route choice behaviour for estimating the multimodal network capacity in a bimodal network. Wang et al. (2020a) applied the combined mode split and traffic assignment equilibrium model to a real-size multimodal network with auto, bus, and metro sub-networks but ignored the transfer modes. All these studies combined mode split and traffic assignment with a fixed origin–destination (OD) demand pattern. However, with the changes in the transport infrastructure, the fixed OD demand assumption may not be valid, as more trips can be generated and distributed to different destinations. To model such a phenomenon, Oppenheim (1993) developed a logit-based function to capture the elastic travel demand, which was then extended by Yang et al. (2000), Ho et al. (2006), and Chu (2011). However, this has not been examined in a multimodal transport network design problem.

The Braess paradox has been in-depth examined in both traffic and transit networks (e.g., Fisk 1979; Hallefjord et al. 1994; Yang and Bell, 1998; Braess et al., 2005; Szeto et al., 2009; Zhao et al., 2014; Szeto and Jiang, 2014b; Jiang et al., 2016; Bagloee et al., 2019, etc.). We refer to Nagurney and Nagurney (2021) for the latest summary of the Braess paradox. Our literature review identifies the following research niches: First, most existing studies focus on a single travel model. Only Fisk (1979) examined a paradox in a bimodal network in which reducing auto flow triggers increased transit costs. Second, very few studies have examined the paradox phenomenon under variable demand. Although Hallefjord et al. (1994) and Bagloee et al. (2019) considered elastic demand, both rely on an explicit demand function to depict the change in demand as a function of travel cost. Third, it has been pointed out that the

1 Walk is considered to access or egress a mode instead of an independent mode to complete a trip.
A paradox can occur in a network with either a flow-dependent or an in-dependent link performance function (Wang and Szeto, 2017; Yao et al., 2019b). None of the existing studies demonstrates the occurrence of a paradox in a network with a mix of link performance functions. In view of the preceding findings, this study is motivated to illustrate the paradox phenomenon in a multimodal transport network under variable demand with different link performance functions.

To conclude, Table 1 provides an overview of the existing literature on bi-level multimodal network design problems and highlights the differences in our model. The contributions of this study are summarised as follows:

1) We developed a bi-level model for the multimodal network design problem to determine the transfer location and capacity simultaneously. As revealed in Table 1, the two decision variables were not considered together. Moreover, both P + R and B + R facilities can be provided in a multimodal transport network.

2) We devised a combined trip distribution/modal split/traffic assignment as the lower-level problem, capturing the demand elasticity and passengers’ route choice behaviour as a result of the construction of the transfer infrastructure. As can be observed from Table 1, the lower-level model in previous studies overlooks the trip-distribution problem. Meanwhile, most of these studies are established on the fixed demand assumption, except Szeto et al. (2015) and Jiang and Szeto (2015), who encapsulated a Lowry-type land-use and transportation interaction model to address the demand elasticity.

3) We conducted experiments to demonstrate the occurrence of a Braess-like paradox phenomenon in a multimodal transport network, stating that providing P + R services may not be beneficial to the system performance measured by the total travel cost, and examine the impacts of different parameters on the occurrence of the paradox.

The remainder of the paper is organised as follows: Section 2 introduces the network representation for the multimodal transport network, assumptions, and notations, and then presents the bi-level formulation. The solution method is described in Section 3. Section 4 presents the experiments to illustrate the occurrence of the Braess-like paradox in a multimodal transport network. Finally, Section 5 presents the conclusions and outlines future directions.

2. Formulation

2.1. Problem description

We consider an urban transport system and adopt a supernetwork (Sheffi, 1985) approach to depict multimodal transportation networks. The supernetwork is denoted by $G = (N, A)$, where $N$ and $A$ represent the set of nodes and links, respectively. It contains $|M|$ subnetworks and each subnetwork is represented by $G^m = (N^m, A^m)$, which corresponds to travel mode $m$, that is, auto, metro or bike. A commuter travelling between nodes $o$ and $d$ can travel via either a single mode in one subnetwork or an intermodal trip through multiple subnetworks, that is, $P + R$ or $B + R$. Accordingly, we use sets $M$ and $\mathcal{M}$ to denote the travel mode that utilises one subnetwork and multiple subnetworks, respectively. For the travel mode $m \in \mathcal{M}$, there exists a set of potential transfer nodes where a traveller can transfer from one subnetwork to another. The link connecting two transfer nodes at different subnetworks is termed a transfer link. The set of all possible transfer links for the mode $m \in \mathcal{M}$ is denoted by $A^m$.

To facilitate the model development, the following assumptions are made.

Table 1

<table>
<thead>
<tr>
<th>References</th>
<th>Upper-level objective</th>
<th>Decision variables</th>
<th>Lower-level model</th>
<th>Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fan et al. (2014b)</td>
<td>Total cost</td>
<td>Location of $P + R$ facilities and transit frequency</td>
<td>Assignment</td>
<td>Fixed</td>
</tr>
<tr>
<td>Szeto et al. (2015)</td>
<td>Sustainability indicators</td>
<td>Link expansion, toll charges, cost recovery</td>
<td>Combined mode split/assignment</td>
<td>Elastic</td>
</tr>
<tr>
<td>Song et al. (2017)</td>
<td>Net social benefit</td>
<td>Location of $P + R$ facilities and transit frequency</td>
<td>Assignment</td>
<td>Fixed</td>
</tr>
<tr>
<td>Liu et al. (2018)</td>
<td>Expected total travel time</td>
<td>Capacity of potential remote $P + R$ facilities</td>
<td>Combined mode split/assignment</td>
<td>Fixed</td>
</tr>
<tr>
<td>Huang et al. (2018)</td>
<td>Operation cost</td>
<td>Transit frequency</td>
<td>Assignment</td>
<td>Fixed</td>
</tr>
<tr>
<td>Wang et al. (2020b)</td>
<td>Total cost</td>
<td>Parking pricing scheme and permits</td>
<td>Assignment</td>
<td>Fixed</td>
</tr>
<tr>
<td>Zheng et al. (2020)</td>
<td>Reserved capacity</td>
<td>Subsidies</td>
<td>Combined mode split/assignment</td>
<td>Fixed</td>
</tr>
<tr>
<td>This study</td>
<td>Total number of trips</td>
<td>Locations of $P + R$ and $B + R$ facilities and capacity</td>
<td>Combined trip distribution/modal split/traffic assignment</td>
<td>Elastic</td>
</tr>
</tbody>
</table>

J. Ye et al.
A1) The link performance function on each subnetwork is independent (e.g., Mesbah et al., 2008; Szeto et al., 2015; Yao et al., 2015). This is reasonable for metro and road subnetworks. As for the car and bus subnetworks, the assumption is valid on the condition that buses are operated in exclusive bus lanes.

A2) In line with the prevailing network design literature (e.g., Beltran et al., 2009; Szeto et al., 2010, 2013; Jiang and Szeto, 2016), the link performance function is continuous and differentiable, and it is formulated as a function of the flow travelling on the link.

A3) The flow on auto subnetwork links refers to vehicle flow, while the flow on metro, bike, and bus links refers to passenger flow. The vehicular flow is transformed into passenger flow by the car occupancy rate (Lam and Small, 2001).

A4) Changes in the network affect residents’ travel behaviours, including trip destination, travel mode, and routes. Meanwhile, it is assumed that commuters’ mode and route choice behaviour follows the multinomial logit distribution in line with the prevailing literature (e.g., Yang et al., 2000; Lam et al., 2008; Liu and Meng, 2014).

A5) For links on each subnetwork, the soft capacity constraint is imposed, meaning that the flow is allowed to be greater than the capacity at an additional cost. For the transfer links, a hard capacity constraint is imposed, representing the limited transfer capacity, such as limited parking spaces in the P + R mode or limited docking slots for the B + R mode.

Based on the preceding assumptions, the multimodal network design problem considered in this study is as follows:

1) In the upper-level, the decision-maker plans to construct infrastructures to enhance urban mobility, measured by the total number of trips generated, and promote the usage of green travel modes, such as P + R and B + R. The corresponding decision variables comprise where to build the infrastructures and their corresponding capacities, that is, at which node in the network the parking space and bike docking slots or public bike share facilities should be constructed under the budget constraint. Mathematically, they are denoted as $\xi_m^a$ and $c_m^a$ respectively. $\xi_m^a$ is a binary variable that equals 1 if candidate link $a$ is selected to construct the transfer infrastructure; otherwise, it equals 0. If transfer link $a$ is selected, the corresponding capacity provided is denoted as $c_m^a$, which is a number between the minimum and maximum capacities, that is, $C_m^a,\min$ and $C_m^a,\max$. Once the capacities associated with transfer links are determined from the supernetwork shown in Fig. 1, the decision-maker can easily infer the transfer nodes at different subnetworks.

As it is already known that providing P + R facilities can generate additional traffic (e.g., Parkhurst, 1995, 2000) and, similarly, linking bike stations and metro stations can lead to increased trips (Noland et al., 2016), the objective of the upper-level model is, thus, devised to examine the maximum number of trips that can be generated, which is bounded by the maximum total travel demand in the area that could be predicted via big data. From such an objective, the decision-maker can infer the maximum number of land-use activities or zonal development potential (Yang et al., 2000).

2) The model in the lower-level depicts the passenger’s behaviour for the newly generated trips. This study simultaneously captures destination choice, travel mode choice, and route choice. Mathematically, we use $q_m^{rs}$ and $q_m^{rs+}$ to denote the number of newly generated trips between OD pairs $rs$ and the number of passengers using mode $m$. Meanwhile, we use $f_m^{p+}$ to represent the number of new passengers travelling via path $p$ of mode $m$. To differentiate between existing and newly generated passengers, we put “0” in the superscript. For example, $f_m^{p0}$ denotes the number of existing passengers travelling via path $p$ of mode $m$, which can be either obtained from the assignment model or estimated using real data.
2.2. Notations

The following notations are used in the proposed bi-level model: The functions of the variables are explained when they are used.

<table>
<thead>
<tr>
<th>Sets</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( M )</td>
<td>set of travel modes utilising a single-mode</td>
<td>set of travel modes utilising multiple modes, i.e., ( P + R )</td>
</tr>
<tr>
<td>( \mathcal{M} )</td>
<td>set of travel modes utilising multiple modes, i.e., ( P + R )</td>
<td>set of travel modes utilising multiple modes, i.e., ( P + R )</td>
</tr>
<tr>
<td>( A^m )</td>
<td>set of links on subnetwork ( m \in M )</td>
<td>set of candidate transfer links for mode ( m \in \mathcal{M} )</td>
</tr>
<tr>
<td>( \overline{A}^m )</td>
<td>set of links on subnetwork ( m \in M )</td>
<td>set of candidate transfer links for mode ( m \in \mathcal{M} )</td>
</tr>
<tr>
<td>( R, S )</td>
<td>set of origins and destinations, respectively</td>
<td>set of origins and destinations, respectively</td>
</tr>
<tr>
<td>( P_{rs} )</td>
<td>set of links connecting OD pair ( rs ) via mode ( m )</td>
<td>set of links connecting OD pair ( rs ) via mode ( m )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Indices</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td>index of links on subnetworks and candidate transfer sites</td>
<td>index of links on subnetworks and candidate transfer sites</td>
</tr>
<tr>
<td>( m )</td>
<td>index of travel mode</td>
<td>index of travel mode</td>
</tr>
<tr>
<td>( p, r, s )</td>
<td>index of path, origin, and destination</td>
<td>index of path, origin, and destination</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{r,s}^m )</td>
<td>maximum demand that can be generated at origin ( r ), destination ( s ), respectively</td>
<td>maximum demand that can be generated at origin ( r ), destination ( s ), respectively</td>
</tr>
<tr>
<td>( \delta_{r}^m )</td>
<td>maximum demand that can be generated at origin ( r ), destination ( s ), respectively</td>
<td>maximum demand that can be generated at origin ( r ), destination ( s ), respectively</td>
</tr>
<tr>
<td>( B )</td>
<td>total budget</td>
<td>total budget</td>
</tr>
<tr>
<td>( C_{r,s}^{\text{max}} )</td>
<td>maximum transfer capacity that could be built at transfer link ( a ) of model ( m )</td>
<td>minimum transfer capacity that should be built at transfer link ( a ) of model ( m )</td>
</tr>
<tr>
<td>( C_{r,s}^{\text{min}} )</td>
<td>minimum transfer capacity that should be built at transfer link ( a ) of model ( m )</td>
<td>minimum transfer capacity that should be built at transfer link ( a ) of model ( m )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decision variables</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower-level decisions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( o_{r,s}^m )</td>
<td>newly generated travel demand between nodes ( r ) and ( s )</td>
<td>newly generated travel demand between nodes ( r ) and ( s )</td>
</tr>
<tr>
<td>( d_{r}^m )</td>
<td>newly generated travel demand of mode ( m ) between nodes ( r ) and ( s )</td>
<td>newly generated travel demand of mode ( m ) between nodes ( r ) and ( s )</td>
</tr>
<tr>
<td>( f_{r}^m )</td>
<td>flow travelling via path ( p ) of mode ( m ) obtained from newly generated demand</td>
<td>flow travelling via path ( p ) of mode ( m ) obtained from newly generated demand</td>
</tr>
<tr>
<td>( v_{r}^m )</td>
<td>flow travelling on link ( a ) of mode ( m )</td>
<td>flow travelling on link ( a ) of mode ( m )</td>
</tr>
<tr>
<td>( \delta_{a}^m )</td>
<td>= 1, if link ( a ) is on path ( p ); otherwise, it equals 0</td>
<td>= 1, if link ( a ) is on path ( p ); otherwise, it equals 0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Upper-level decisions</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( e_{a}^m )</td>
<td>designed capacity for travel mode ( m ) at transfer link ( a )</td>
<td>designed capacity for travel mode ( m ) at transfer link ( a )</td>
</tr>
<tr>
<td>( e_{a}^m )</td>
<td>= 1, if transfer link ( a ) is selected; otherwise, it equals 0</td>
<td>= 1, if transfer link ( a ) is selected; otherwise, it equals 0</td>
</tr>
</tbody>
</table>

2.3. Bi-level formulation

The bi-level optimisation model developed for the multimodal network design problem is presented in the following two subsections.

2.3.1. Upper-Level problem

\[
\max_{\xi, p, d} \sum_{r \in R} o_{r}^+ = \sum_{r \in R} o_{r}^+
\]

subject to

\[
G(\xi) \leq B \tag{2}
\]

\[
o_{r}^+ = \sum_{a \in A^m} \sum_{o \in M} q_{a,o}^m (o) \leq o_{r}^{\text{max}} - o_{r}^+, \forall r \in R \tag{3}
\]

\[
d_{r}^+ = \sum_{p \in R} \sum_{m \in M} q_{a,p}^m (o) \leq d_{r}^{\text{max}} - d_{r}^+, \forall s \in S \tag{4}
\]

\[
\varepsilon_{a}^m c_{a}^{\text{min}} \leq v_{a}^m \leq \varepsilon_{a}^m c_{a}^{\text{max}}, \forall a \in \overline{A}^m, m \in \mathcal{M} \tag{5}
\]

\[
v_{a}^m (o) \leq \varepsilon_{a}^m, \forall a \in \overline{A}^m, m \in \mathcal{M} \tag{6}
\]

\[
\varepsilon_{a}^m \in \{0, 1\}, \forall a \in \overline{A}^m, m \in \mathcal{M} \tag{7}
\]

\[
o_{r}^+ \geq 0, \forall r \in R \tag{8}
\]

\[
d_{r}^+ \geq 0, \forall s \in S \tag{9}
\]

The objective of the upper-level problem is to maximise the total number of trips generated. \( \xi = [\xi_a^m] \), \( c = [c_a^m] \), \( o = [o_r^+] \), and \( d = [d_s^+] \) are the vector notations for the corresponding decision variables. Eq. (2) is the budget constraint, where \( G(\xi) \) is the cost function of constructing infrastructures and is generally assumed to be non-negative, increasing, and differentiable (Yang and Bell, 1998) with
respect to capacity. Eqs. (3) and (4), respectively, set the upper and lower bounds for the number of trips generated at each origin and destination node. \( g^n_a(a) \) denotes the newly generated trip travelling via mode \( m \) between nodes \( r \) and \( s \). It will be determined at the lower-level once the total number of new trips is specified. Therefore, it can be expressed as a function of \( a \). Eq. (5) represents the capacity constraints associated with each building of transfer infrastructure. If a transfer location is determined, that is, \( c^n_{wh} = 1 \), then the capacity to build should be within \([C^n_{min}, C^n_{max}]\); otherwise, \( c^n_{wh} = 1 \) and \( c^n_{wh} = 0 \). Eq. (6) is the capacity constraint at the transfer link, stating that the flow traversing link \( a \), denoted by \( \phi^n_{wh}(a) \), should be no greater than the link capacity. The capacity constraint applies to the travel mode \( m \in \mathcal{M} \), in which the transfer flow is strictly bounded by the infrastructure capacity, that is, the number of parking slots. Eq. (7) is the definitional constraint for the variable \( c^n_{wh} \). Finally, Eqs. (8) and (9) are nonnegative constraints for variables \( o^n_r \) and \( d^n_r \), respectively.

### 2.3.2. Lower-level problem

At the lower-level, a combined trip distribution/modal split/traffic assignment model was developed.

\[
\min_{q, t} Z = z_1 + z_2 + z_3 + z_4
\]

where

\[
z_1 = \frac{1}{\theta} \sum_{m \in \mathcal{M}, \mathcal{M}} \int_{P_{rs}} f_\mu (\ln q_\mu^m - 1) + \sum_{m \in \mathcal{M}, \mathcal{M}} \sum_{a \in \mathcal{A}} \int_0^{\phi_a^m} \ell_a(x) dx
\]

\[
z_2 = \sum_{r \in \mathcal{R}, s \in \mathcal{S}} \sum_{m \in \mathcal{M}, \mathcal{M}} f_m^m (\ln q_m^m - 1)
\]

\[
z_3 = \frac{1}{\theta} \sum_{r \in \mathcal{R}, s \in \mathcal{S}} \delta^r (\ln q^r - 1)
\]

\[
z_4 = -\int_0^{\sum_{m \in \mathcal{M}} h_m} h_m(x) dx
\]

subject to

\[
g^n_{rs} = \sum_{\mu \in \mathcal{M}} f_{\mu}^m, \forall r \in \mathcal{R}, s \in \mathcal{S}, m \in \mathcal{M} \cup \mathcal{M}
\]

\[
g^n_{rs} = \sum_{m \in \mathcal{M}, \mathcal{M}} f_{\mu}^m, \forall r \in \mathcal{R}, s \in \mathcal{S}
\]

\[
f_{\mu}^m = f_{\mu}^m + f_{\mu}^m, \forall m \in \mathcal{M}, \mathcal{M}, \mu \in \bigcup_{r \in \mathcal{R}, s \in \mathcal{S}} P_{rs}^m
\]

\[
o^n_r = \sum_{m \in \mathcal{M}} g^n_{rs}, \forall r \in \mathcal{R}
\]

\[
v^n_m = \sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} \sum_{m \in \mathcal{M}, \mathcal{M}} \left( f_{\mu}^m + f_{\mu}^m \right) \delta_{rs}, \forall a \in \mathcal{A} \cup \mathcal{A}, m \in \mathcal{M} \cup \mathcal{M}
\]

\[
f_{\mu}^m, f_{\mu}^m \geq 0, \forall m \in \mathcal{M} \cup \mathcal{M}, \mu \in \bigcup_{r \in \mathcal{R}, s \in \mathcal{S}} P_{rs}^m
\]

\[
g^n_{rs}, q^n_{rs} \geq 0, \forall r, s \in \mathcal{S}, m \in \mathcal{M} \cup \mathcal{M}
\]

The objective function of the lower-level model contains four elements defined by Eqs. (11)–(14). \( z_2 \) replicates the mathematical formulation for the stochastic user equilibrium. \( z_2 \) and \( z_2 \) represent the entropy functions on mode choice and destination choice behaviour, respectively. \( z_4 \) is the inverse of the elastic demand function. Eqs. (15), (16), and (18) are the flow conservation constraints. Eq. (17) states that the number of passengers travelling via path \( \mu \) of mode \( m \) includes the newly generated passengers who choose the path in question and the existing passengers who use this path. Eq. (19) is the definitional constraint for link flow and path flow. For consistency, it is assumed that the modal and route choices of existing commuters are characterised via the multinomial logit discrete choice model, which is the same as the newly generated travel demand. By examining the Karush–Kuhn–Tucker (KKT) conditions of the lower-level problem, it can be proved in the Appendix that the demand distribution, travellers’ mode choice, and route choice follow a multinomial logit distribution.

Compared with the model developed by Yang et al. (2000), our contributions are twofold. First, the stochasticity in passengers’ route choice is considered in the sense that passengers’ route choice behaviour follows the stochastic user equilibrium instead of user
equilibrium. Second, an additional mode choice component is encapsulated to capture travellers’ mode choice behaviour and prove that the resultant mode choice follows a multinomial logit model.

2.3.3. Properties
Here, we briefly discuss the following properties of the bi-level model.

1) Solution existence. The existence of a solution to the lower-level problem can be generally guaranteed because the link performance functions are continuous and differentiable, and the solution space is bounded by the travel demand (Sheffi, 1985). Thus, the feasibility of the bi-level model depends on whether the capacity constraint in the upper-level is violated. Because the hard capacity is only applicable to the newly constructed transfer infrastructure, a feasible solution that always exists is the doing nothing solution.

2) Capacity constraint. In the bi-level formulation, the flow capacity constraint can be placed in either the upper- or lower-level model. Nevertheless, as noted in Yang et al. (2000), there are different behavioural implications, and the results vary. When it is placed at the lower-level, the resultant model yields a queuing time on saturated links. The Lagrange multipliers associated with binding capacity constraints are normally interpreted as queuing delay times (e.g., Bell, 1995; Lam et al. 1999, 2002). In such a case, the number of newly generated trips tends to reach its maximum value at the expense of the queuing time. In contrast, when the capacity constraint is imposed at the upper-level, it leads to an assignment model with no queues. Accordingly, the number of newly generated trips may not reach the maximum value, depending on whether there exists a saturated link. Finally, it is not necessary to put it at both levels, as in such a case, the capacity constraint in the lower-level will ensure that the capacity constraint is satisfied in the upper-level automatically.

3. Solution algorithm
To solve the bi-level programming problem, most existing studies adopt heuristic (Yang, 1995; Chiou, 2005; Angelo and Barbosa, 2015), metaheuristic (Fan et al., 2014a; Koh, 2007), and matheuristic (Szeto and Jiang, 2012; Jiang et al. 2013; Szeto and Jiang, 2014a; Carosi et al., 2019; Liu et al., 2019; Jiang, 2021) methods, given that a bi-level network design problem is a well-known NP-hard problem. Because the primary purpose of this study is to examine the effect of optimising transfer capacity under variable demand, the classic GA is employed to solve the bi-level model.

It should be noted that although our problem exhibits a format similar to that in Yang et al. (2000), the successive linear programming (SLP) in their study cannot be directly applied to solve our bi-level model because our upper-level model is formulated as a mixed-integer linear programming problem. Nevertheless, once the decision variables $\xi$ and $\rho$ are given, the resultant upper-level model is reduced to a linear programming method, and the SLP method proposed by Yang et al. (2000) can be adopted. Therefore, in the GA, only $\xi$ and $\rho$ are encoded as decision variables.

In our preliminary experiments, it was found that a direct application of the SLP algorithm depicted in Yang et al. (2000) may not guarantee the convergence of the model because the step size for updating the decision variables in each iteration is insufficiently described in their study. Therefore, inspired by Zhang et al. (1985) and Ferris and Zavriev (1996), this study develops an adaptive step size method for updating a trust region, where the trust region can be understood as the lower and upper bounds of the decision variables in the linear programming model, and the adaptive step size determines the changes in the size of the trust region. The algorithm is described below, where the superscript $k$ is introduced to indicate the iteration number when necessary.

**Algorithm:** An SLP with an adaptive step size for solving the maximum trip generation model.

Step 1: Initialise the iteration counter $k = 0$, set the initial lower and upper bounds for $[\alpha_r^k]_{r \in R}$ as $[\alpha_{r,\min}^k, \alpha_{r,\max}^k]_{r \in R}$, set parameters for adjusting the trust region $\rho_r \in (0, 1), \forall r \in R$, and $[\rho_r^k]_{r \in R}$, and set $\varepsilon$ as the parameter governing the convergence of the algorithm.

Step 2: Determine initial trip distribution pattern $q^k = [\alpha_r^k]_{r \in R}$.

Step 3: Solve the lower-level combined trip distribution/modal split/traffic assignment model and obtain $[q_{rs}^{m,k}]$ and $[f_p^{m,k}]$.

Step 4: Calculate the approximated influence factors $l_{ur}$ and $l_{ur}$.

Step 5: Formulate a linear approximation model for the upper-level model by replacing $q_{rs}^k(o)$ in Eqs. (3) and (4) and $v_a(o)$ in Eq. (6).

Step 6: Solve the resultant linear program and obtain $\sigma^{k+1} = [\alpha_r^{k+1}]_{r \in R}$.

Step 7: Check convergence.

If $\max_{r \in R} |\alpha_r^{k+1} - \alpha_r^k| / \alpha_r^k \leq \varepsilon$, then stop; otherwise, proceed to the next step.

Step 8: Update the step size.

If $|\alpha_r^{k+1} - \alpha_r^k| \leq \rho_r^k$, then reduce the trust region by setting $\alpha_{r,\min}^{k+1} = \rho_r \cdot \alpha_{r,\min}^k$, $\alpha_{r,\max}^{k+1} = \rho_r \cdot \alpha_{r,\max}^k$, and $\rho_r^{k+1} = \rho_r^k \cdot \rho_r$.

Step 9: Set $k = k + 1$ and return to Step 3.

Remarks of the above algorithm. (1) The maximum trip generation model is a bi-level model with given $\xi$ and $\rho$. (2) In Step 1, the initial solution region, that is $[\alpha_{r,\min}^k, \alpha_{r,\max}^k]$, for the decision variables can be set large, as it will shrink over iterations. The notation
with subscript $r$ implies that each decision variable $o^r_j$ is associated with its parameters for updating the step size. $\rho_r$ does not contain the superscript $k$, as the algorithm adopts a constant rate to update the solution region. In the future, it will be possible to develop a scheme that can update the ratio over iterations. (3) In Step 3, the influencing factors can also be obtained via a sensitivity analysis (e. g., Tobin and Friesz, 1988; Yang, 1997). (4) Compared with the method proposed by Yang et al. (2000), our extensions are twofold. One is to include a trust region when solving the linear programming problem, and the other is to update the trust region adaptively in Step 8. (5) The proof of convergence when using the SLP method to solve the nonlinear optimisation programming problem can be found in Zhang et al. (1985) and Ferris and Zavriev (1996). However, as our model is a bi-level model, a theoretical proof of convergence and global optimality may not be easily established. It can be a challenging research direction to be explored in the future. In this study, we only examined the convergence numerically.

4. Numerical study

The numerical studies were designed to illustrate the properties of the model and examine two Braess-like paradox phenomena in a multimodal transport network. The algorithm was coded using MATLAB 2018. All the input data and results were deposited in the author’s GitHub repository (https://github.com/hkujy/CapacitedMNDP).

4.1. Braess-like paradox under fixed demand

4.1.1. Occurrence of the paradox

The first experiment was conducted to demonstrate that introducing P + R services could induce a Braess-like paradox in a multimodal transport network under a fixed OD demand. We constructed a small network that resembled the classic Braess network, as shown in Fig. 2. The network contains four nodes, where node B represents a metro station and node A denotes a potential parking area close to node B. There are four links, and their travel modes and link performance functions are listed in the figure. The travel demand between nodes O and D was set to 2000. Compared with the original setting in the Braess network, our network contains links with a constant travel time to resemble travellers’ walking time. Fig. 2(a) and (b), respectively, represent the network before and after constructing parking slots at node A, so that P + R is available. In the previous scenario, there are two travel modes, and each mode has only one path, as shown in Fig. 2(a), that is, using a private vehicle via path O-A-D and using a metro via path O-B-D. In the following scenario, when a parking area is constructed at node A, the P + R mode is available to the travellers via paths O-A-B-D. To obtain the flow distribution in the network, only the lower-level model is solved because the upper-level objective is constant under a fixed demand. Accordingly, a capacity constraint is imposed on the lower-level problem.

Table 2 summarises the total travel time and flow distribution obtained before and after the scenarios. This shows that the total travel time in the after scenario is higher than that in the before scenario. This is similar to the Braess paradox (Braess, 1968), which states that adding a new link deteriorates the network performance measured by the total travel time due to travellers’ selfish routing behaviour. In our example, this is because the travel time of the P + R mode (path O-A-B-D) is much less than that of the other two modes; thus, it attracts most of the OD travellers. On the one hand, the travel time via link 2 slightly decreases because of the reduction in the number of car users. On the other hand, the travel time associated with link 1, through with travellers’ access to parking facilities, increases significantly. Thus, the overall travel time of the network increases, and a paradox occurs. Furthermore, from a mathematical point of view, the formulation of the lower-level problem only ensures that the assignment results follow a logit distribution instead of distributing passengers to achieve system optimisation.

### Link performance functions

\[
t_1^{\text{car}} = 4 + \left( \frac{v_1^{\text{car}}}{500} \right)^4, \quad t_2^{\text{car}} = 43 + \left( \frac{v_2^{\text{car}}}{1000} \right)^4, \quad t_3^{\text{walk}} = 25, \quad t_4^{\text{metro}} = 25 + \frac{v_4^{\text{metro}}}{500}
\]

**Table 2** summarises the total travel time and flow distribution obtained before and after the scenarios.

<table>
<thead>
<tr>
<th>Path</th>
<th>O - A - D (Car)</th>
<th>O - B - D (Metro)</th>
<th>O - A - B - D (P + R)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(a) Before building car park at node A</td>
<td>(b) After building car park at node A</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Network for illustrating the Braess-like paradox phenomena.
4.1.2. Effect of capacity on the occurrence of the paradox

This section examines the effect of capacity constraints, that is, limited parking space, on the occurrence of the paradox, which has been ignored in previous examples by assuming that there is enough parking space. In the test, the capacity varied from 100 to 2000. Fig. 3(a) shows that the total travel time starts to decline and reaches a stable level after the capacity is greater than 950. Fig. 3(b) shows that the capacity is binding despite the value of \( \gamma \), as the flow on path 3 equals the capacity all the time. In contrast, in Fig. 4 (c) and (d), the capacity constraint is not binding initially, which explains the flow associated with path 3. Meanwhile, it is observed that the flow travelling via path 3 always equals the capacity when the capacity is set below 950, meaning that the capacity constraint is binding. This implies that the number of travellers willing to travel via the P + R could be larger than the available capacity, which could induce a queuing delay time (see Section 2.3.3 for the discussion of the model properties) added to the travel time of path 3. The results justify the necessity of optimising the capacity of parking space for P + R services in terms of reducing the total travel time.

4.1.3. Effect of the scale parameters on the occurrence of the paradox

This section illustrates the effect of the scaling parameter in the logit model on the occurrence of the paradox. Theoretically, when its value increases to infinity, the result of the stochastic assignment model is close to that obtained from the user equilibrium model. In this experiment, the scale parameter was varied from 0.05 to 0.55 and after scenarios but does not eliminate it. Fig. 4 (a) shows that a higher value of the scale parameter diminishes the travel time gap between the before and after scenarios, whereas Fig. 4 (b) displays the corresponding changes in the path flow.

Fig. 3(a) shows that a paradox occurs as long as the parking space is provided. The total travel time reached the maximum when its capacity was 650. Afterwards, the total travel time starts to decline and reaches a stable level after the capacity is greater than 950. Fig. 3(b) shows that the increase in parking capacity attracts more travellers to use the P + R model, represented by the climbing of the flow associated with path 3. Meanwhile, it is observed that the flow travelling via path 3 always equals the capacity when the capacity is set below 950, meaning that the capacity constraint is binding. This implies that the number of travellers willing to travel via the P + R could be larger than the available capacity, which could induce a queuing delay time (see Section 2.3.3 for the discussion of the model properties) added to the travel time of path 3. The results justify the necessity of optimising the capacity of parking space for P + R services in terms of reducing the total travel time.

4.2. Paradox under variable demand

In the following experiments, we adopted the network shown in Fig. 2 and varied the parking capacity from 100 to 2000. The existing and maximum demands were set to \( o^0 = 100 \) and \( o^{\text{max}} = 2000 \), respectively. The results are presented in Fig. 5.

Fig. 5(a) and (b), respectively, plot the changes in the flow of the three paths and the number of newly generated trips under variable capacities. In general, it can be observed that both the path flow and the number of newly generated trips increase with the transfer capacity until the number of newly generated trips reaches the maximum total demand. Notably, the flow travelling via path 2 grows more dramatically compared with those travelling via paths 1 and 3. Our explanation is that the changes in the travel time of path 1 are much gentler with the increase in parking capacity and plotted the results in Fig. 5(a) that the curves corresponding to \( \tau = 600 \) and \( \tau = 1000 \) start from the same point and have equal total travel time.

Under variable demand, the total travel time may not be a consistent measure because of changes in travel demand. Therefore, we examined the proportion of travellers using P + R with the increase in parking capacity and plotted the results in Fig. 5(d). An interesting phenomenon is observed: expanding parking capacity could reduce the proportion of travellers using P + R services. This

![Changes in total travel time](image1)

(a) Changes in total travel time

![Changes in path flow](image2)

(b) Changes in path flow

Fig. 3. Effect of the transfer capacity on the occurrence of the Braess-like paradox.
indicates that adding more parking spaces does not make the P+R services more attractive to the newly generated travel demand. This could be considered a paradox because it is in contrast with our intuition that increasing the P+R capacity would increase the attractiveness of such a service. Nevertheless, despite the decrease in the share of P+R services, the market share of metro services continues to grow mildly. In terms of the path cost, as expected, it increases with the improvement in transfer capacity because a higher transfer capacity permits more travel demand to be generated in the network and induces higher congestion costs.

4.3. Discussion of paradox

In summary, experiments under fixed demand extend the Braess paradox to a multimodal context. Compared with existing studies, our network contains both flow-dependent and non-flow-dependent link performance functions to capture passengers’ travel times via different modes and transfer walking time. Nevertheless, we acknowledge that the network and link performance are hypothetical, and the results are theoretical. Thus, future efforts are required to explore the empirical evidence. As the occurrence of the Braess paradox has been noted in real-world road networks (Bagloee et al., 2019), it might not be impossible to detect it in a multimodal transport network.

Under elastic demand, our findings concur with those in the existing literature (Parkhurst, 1995, 2000; Noland et al., 2016) in the sense that providing parking and biking stations induces more trips. Interestingly, our case study in Section 4.2 shows that, although the total number of trips for P+R continues to increase, a high proportion of the newly generated travel demand travels via the public transport model. Intuitively, this is because providing parking space generates a number of trips, whereas the number is larger than the parking capacity. Then, passengers cannot travel via the P+R and are distributed between public and private transport. Because the travel cost associated with public transport (e.g., metro in the example) grows mildly with the increase in travel demand compared with the cost associated with private transport, more passengers select public transport. The findings of this experiment underpin the necessity of maintaining a high quality of public transport when providing P+R services.

4.4. Effect of multimodal transfer capacity

This experiment was conducted to illustrate the design of the optimal transfer capacity under budget constraints in a multimodal transport network. We consider a more general network containing 12 links and seven nodes, as shown in Fig. 6. There are two transfer
infrastructure options: one is to build parking spaces at node 7 to facilitate P + R trips, and the other is to construct bike docking slots at node 5 to promote B + R users. There are two OD pairs in this network, that is, OD pair (1, 4) and OD pair (6, 4), with the same existing demand of 100 commuters. Detailed route data and link performance functions can be assessed from the Github repository of the author.

4.4.1. Convergence of the SLP with an adaptive step size

Before demonstrating the effects of the capacity constraint, we first examine the performance of the SLP algorithm with an adaptive step size. The parameter for regulating the step size is increased from 0.3 to 0.9. In the experiment, the capacities associated with links 5–2 and 7–3 were set to 100, and the scale parameters were set at 0.1. The convergent plots are presented in Fig. 7. It can be observed that the algorithm converges under the proposed adaptive step size scheme.

The solutions obtained for different step sizes are listed in Table 3. First, it can be observed that the flow travelling via link 7–3 approaches the link capacity. The binding of this link confines the number of newly generated travel demands. There is even a residual capacity on link 5–2. No additional travel demand will be generated. Second, the solution obtained under $\rho_r = 0.3$ deviates from that obtained when $\rho_r = 0.6$ and $\rho_r = 0.9$. This is because a smaller step size may cause the algorithm to converge to a local optimal solution.
4.4.2. Effect of capacity constraint

In this experiment, the capacities of the two links simultaneously varied from 0 to 400. The total newly generated demands under different scale parameters are illustrated in Fig. 8, where the x- and y-axis denote the capacity of links 5–2 and link 7–3, respectively, and the contour represents the new trips generated.

The following conclusions can be drawn.

1) First, it is obvious that increasing the capacity attracts more travel demand. This reveals that providing more infrastructure could lift the total travel demand.

2) Interestingly, expanding only one link could induce a higher travel demand than expanding both. For example, in Fig. 8(a), the number of newly generated trips is higher when the capacity of link 5–2 is 100 than when both capacities are 50. This implies that when the budget is tight, it is better to provide sufficient transfer capacity at one location. This rationalizes one of the motivations of this study. In other words, we incorporate the decisions of where to construct the transfer infrastructure into a multimodal transport network design problem.

3) When the transfer infrastructure is constructed at both links, expanding the capacity of one link may not further induce higher demand. For example, in Fig. 8(a), when the capacity of link 7–3 equals 50, the total number of newly generated demand is no more than 400 and remains constant despite the growth in the capacity of link 5–2. This is because, as revealed in Table 3, the capacity of one of the links is binding, which restricts the growth of the total travel demand. This means that the capacity must be increased

**Table 3**

<table>
<thead>
<tr>
<th>Value of $\rho_r$</th>
<th>Link 5–2</th>
<th>Link 7–3</th>
<th>Newly Generated Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$v_a$</td>
<td>$v_a/c_a$</td>
<td>$v_a$</td>
</tr>
<tr>
<td>0.3</td>
<td>90.76</td>
<td>0.91</td>
<td>99.98</td>
</tr>
<tr>
<td>0.6</td>
<td>90.81</td>
<td>0.91</td>
<td>100.00</td>
</tr>
<tr>
<td>0.9</td>
<td>90.80</td>
<td>0.91</td>
<td>99.98</td>
</tr>
</tbody>
</table>

*Note: $v_a$ is the link flow and $v_a/c_a$ is the flow/capacity ratio.

**Fig. 7.** Convergence of the SLP with an adaptive step size.
simultaneously, and the capacity at each location should be optimised. This justifies the necessity of developing an optimisation model to determine the capacity allocation at different stations.

4) Finally, as can be observed by comparing the two subfigures, more demand is generated at a lower scale parameter. This is because the path flows are more spread among different options when the scale parameter is low. Accordingly, the path that traverses the capacitated links attracts fewer passengers, and these links become binding at a higher number of newly generated demand.

4.5. Optimal design in the Sioux-Falls network

The Sioux-Falls network is adopted for investigating the optimal design of multimodal transfer nodes using GA. Since the original Sioux-Falls network only contains auto links, this study adds a bike network as well as the transit lines to the Sioux-Falls network to mimic a multimodal urban transport network (see Fig. 9). There are total 65 nodes and 121 links in the modified multimodal Sioux-Falls network. The origin nodes 3, 2, 13, 20, 14 are respectively identical with the nodes 26, 36, 35, 41, 32 and the nodes 43, 57, 45, 63, 48 in physical space while the destination nodes 9, 10, 8, 16 are identical with the nodes 29, 30, 38, 39 and the nodes 52, 53, 59, 16 in physical space. The maximum potential demand of each origin is 2000. The central urban area is highlighted by the blue oval. The 8 green nodes and dotted links are the candidate transfer links. The total construction cost function is given by

\[ G = 5 \times c_\text{P+R}_{\min} + c_\text{B+R}_{\max} \]

and the total budget 300. More detailed link attributes and the demand setting can be obtained in the author’s GitHub.

To solve bi-level model, we use the Matlab GA toolbox with the proposed SLP algorithm. Meanwhile, the GA was run for 8 times with different random seeds. The algorithm took 45242.9 s to finish 100 iterations on average, and the average best objective value over 100 iterations is plotted in Fig. 10. The best objective value of the 8 runs is 4275. In the optimal solution, there are no parking spaces suggested to be constructed at P + R transfer nodes, but 300 biking docks constructed at node 62.

The output results show that if the optimal objective is the network capacity, the B + R transfer node is preferred as it occupies a smaller area per unit. It is worthy of acknowledging the results of the model represent an ideal scenario under various assumptions. Particularly, the premise that drivers will switch their travel mode from P + R to B + R as long as the parking infrastructure is replaced with biking infrastructure might be strong and unrealistic. Nevertheless, the results imply the potential of promoting the usage of the bike to enhance urban mobility. Therefore, it will encourage the decision-maker to consider and implement other strategies that could achieve a joint effect in promoting the usage of bicycles, such as better integration of bike and public transport, expanding and construing dedicated bike lanes, and providing a safe and healthy cycling environment, etc.

5. Conclusion

In view of the boom in multimodal mobility, this study develops a bi-level model for the multimodal network design problem. The upper-level problem is to determine the location and capacity of the transfer infrastructure to be built simultaneously. The lower-level problem is the assignment of the combined trip distribution/modal split/traffic. It is proved that at optimality, the trip distribution and commuters’ route-choice behaviour depicted by the optimal solution of the lower-level problem follow a multinomial logit

Fig. 8. Effect of capacity constraint on the total newly generated demand.
distribution. Numerical studies were conducted to illustrate the optimal design of transfer capacity in a multimodal transport network and a Braess-like paradox phenomenon in a multimodal transport network, stating that constructing parking space to stimulate the usage of $P + R$ service may induce higher total passenger travel time. Moreover, it is found that under variable demand, the modal share of the $P + R$ service for the newly generated travel demand could decline with an increase in the capacity of parking spaces.

This study provides various research directions. For example, (1) the lower-level model assumes that the travellers’ route choice follows the multinomial logit model. It would be more interesting to examine other behavioural models, such as nest-logit (e.g., Zheng

Fig. 9. Modified multimodal Sioux-Falls network.
et al., 2020) or path-size logit (e.g., Liu et al., 2019). (2) The network and link performance functions used in the example to demonstrate the occurrence of the paradox are hypothetical for theoretical interests. In practice, as the Braess paradox has been observed in road networks in the real world (Bagloee et al., 2019), it is expected that such a phenomenon can also be observed in a multimodal transport network. Other paradoxes like the Downs-Thomson paradox, which were not considered in this study, might also occur in a multimodal transport network. (Ku et al., 2020; Prieto Curiel et al., 2021; Wang et al., 2021a). Thus, one future research direction would be to use empirical studies to detect different paradoxes in a multimodal transport network. (3) Once the paradox is observed, it is crucial to develop a method to avoid it. Bazzan and Klügl (2005) noted that this could be avoided by providing route recommendations. With the emerging personalised mobility, we propose exploring personalised route guidance (e.g., Ceder and Jiang, 2020; Jiang and Ceder, 2021). (4) The lower-level developed in this study relies on generating a path choice set in advance. A large path set may impede the application for a large network and the choice set could affect the solution, as discussed in the traditional stochastic user equilibrium model (e.g., Prashker and Bekhor, 2004; Bekhor and Toledo, 2005). Developing a link-based or approach-based assignment model would be a future research avenue to overcome these limitations (e.g., Szeto and Jiang, 2014a; Jiang and Szeto, 2016; Jiang, 2021). Meanwhile, it is also possible to consider the land use effect in lower-level model (Zhong et al., 2021a, 2021b). (5) Regarding demand elasticity and trip distribution, this study adopts a simplified function to represent the destination’s attractiveness. One way to improve this would be to incorporate more advanced land-use and transportation interaction models (e.g., Zhong et al., 2021a, 2021b).

CRediT authorship contribution statement

Jiao Ye: Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. Yu Jiang: Conceptualization, Methodology, Validation, Writing – review & editing, Formal analysis, Supervision. Jun Chen: Conceptualization, Funding acquisition. Zhiyuan Liu: Conceptualization, Supervision. Renzhong Guo: Resources, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors thank three anonymous reviewers for their valuable suggestions and constructive comments for improving the quality of the paper. This research is funded by the Scholarship Fund of China Scholarship Council (CSC No. 201806090216), the National Natural Science Foundation Council of China under Project No. 51638004 and No. 71971038, the Innovation Fund Denmark (IFD) under File No. 4109-00005.
Appendix

This Appendix shows that, at optimality, demand distribution, travellers’ mode choice, and route choice obtained by solving the lower-level problem follows a multinomial logit distribution. This can be proved by examining the KKT conditions of the lower-level problem. First, we provide the Lagrange function for the lower-level problem.

\[ L = Z + \sum_{r \in R} \sum_{x \in X} \sum_{m \in M} \alpha_{r}^{m} \left( q_{r}^{m} - \sum_{p \in P_{r}} f_{p}^{m} \right) + \sum_{r \in R} \sum_{x \in X} \sum_{m \in M} \beta_{r}^{m} \left( q_{r}^{m} - \sum_{m \in M} f_{r}^{m} \right) + \left( \sum_{m \in M} \sum_{p \in P_{r}} \eta_{p}^{m} \left[ f_{p}^{m} - \left( f_{p}^{m+} + f_{p}^{m-} \right) \right] \right) \]  

where \( \alpha_{r}^{m}, \beta_{r}^{m}, \) and \( \eta_{p}^{m} \) are the dual variables associated with Eqs. (15)–(17). Constraints (18) and (19) are directly substituted in Eqs. (13) and (11), and they are captured in objective function \( Z \).

Travellers’ route choice follows multinomial logit distribution

To prove this, we consider the KKT conditions with respect to the variable \( f_{p}^{m} \). At optimality, the following condition holds:

\[ f_{p}^{m} \delta_{p} = 0 \text{ and } \frac{\partial L}{\partial f_{p}^{m}} \geq 0, \quad \forall p \in P_{r}, m \in M \cup M \]  

where

\[
\frac{\partial L}{\partial f_{p}^{m}} = \frac{\partial Z}{\partial f_{p}^{m}} + \frac{\partial}{\partial f_{p}^{m}} \left( \sum_{r \in R} \sum_{x \in X} \sum_{m \in M} \alpha_{r}^{m} \left( q_{r}^{m} - \sum_{p \in P_{r}} f_{p}^{m} \right) \right) + \frac{\partial}{\partial f_{p}^{m}} \left( \sum_{r \in R} \sum_{x \in X} \sum_{m \in M} \beta_{r}^{m} \left( q_{r}^{m} - \sum_{m \in M} f_{r}^{m} \right) \right) + \frac{\partial}{\partial f_{p}^{m}} \left( \sum_{m \in M} \sum_{p \in P_{r}} \eta_{p}^{m} \left[ f_{p}^{m} - \left( f_{p}^{m+} + f_{p}^{m-} \right) \right] \right) = \frac{\partial(\Delta_{p} + \Delta_{m} - \Delta_{p}^{m})}{\partial f_{p}^{m}} - \alpha_{r}^{m} \Delta_{p}^{m} - \eta_{p}^{m}
\]

where \( \Delta_{p} = 1 \) if path \( p \) of mode \( m \) connects OD pair \( rs \); otherwise, \( \Delta_{p}^{m} = 0 \).

By substituting Eqs. (11)–(14) into Eq. (A.3), we obtain

\[
\frac{\partial L}{\partial f_{p}^{m}} = \frac{\partial}{\partial f_{p}^{m}} \left( \frac{1}{\theta} \sum_{m \in M} \sum_{p \in P_{r}} f_{p}^{m} \left( \ln f_{p}^{m} - 1 \right) + \sum_{m \in M} \sum_{x \in X} \int_{0}^{r} t_{a}(x)dx \right) = \frac{1}{\theta} \left( \ln f_{p}^{m} - 1 \right) + f_{p}^{m} \left( \frac{1}{f_{p}^{m+}} \right) + \sum_{m \in M} \sum_{x \in X} t_{a}(v_{a}^{m}) \delta_{ap} - \alpha_{r}^{m} \Delta_{p}^{m} - \eta_{p}^{m}
\]

Then, when \( f_{p}^{m} > 0 \), the KKT conditions imply that

\[
\frac{1}{\theta} \ln f_{p}^{m} + \sum_{m \in M} \sum_{x \in X} t_{a}(v_{a}^{m}) \delta_{ap} - \alpha_{r}^{m} \Delta_{p}^{m} - \eta_{p}^{m} = 0
\]

\[
\Rightarrow f_{p}^{m} = \exp \left( -\theta \left( \sum_{m \in M} \sum_{x \in X} t_{a}(v_{a}^{m}) \delta_{ap} - \alpha_{r}^{m} \Delta_{p}^{m} - \eta_{p}^{m} \right) \right)
\]

Consider another path \( p' \) with \( f_{p'}^{m} > 0 \), and the above condition also holds. Therefore, we derive the following equation:
\[ f_p^+ = \exp \left( -\theta \left( \sum_{m \in M} \Delta^m \delta^m - \alpha^m \Delta^m - \eta^m \right) \right) \]
\[ f_p^- = \exp \left( -\theta \left( \sum_{m \in M} \Delta^m \delta^m - \alpha^m \Delta^m - \eta^m \right) \right) \]
which implies that the passengers’ route choice is in line with the multinomial logit distribution.

**Travellers’ mode choice follows multinomial logit distribution**

We consider the KKT conditions with respect to variable \( q_r^m \), which is

\[ \frac{dL}{dq_r^m} = 0 \text{ and } \frac{dL}{dq_r^m} \leq 0, \quad \forall p \in P_r^m, \quad m \in M \cup \bar{M} \]

where

\[
\frac{dL}{dq_r^m} = \frac{dZ}{dq_r^m} + \sum_{r \in R} \sum_{m \in M} \sum_{r \in S} \alpha^m \left( q_r^m - \sum_{p \in P_r^m} f_p^m \right) + \frac{\partial}{\partial q_r^m} \left( \sum_{m \in M} \sum_{p \in P_r^m} \eta^m \left[ f_p^m - \left( f_p^m + f_p^m \right) \right] \right) = \frac{\partial (z_1 + z_2 + z_3 + z_4)}{\partial q_r^m} + \alpha^m - \beta^m
\]

By substituting Eqs. (11)–(14) into (A.8), we obtain

\[
\frac{dL}{dq_r^m} = \frac{1}{\gamma} \sum_{r \in R} \sum_{m \in M} \sum_{r \in S} \alpha^m \left( \ln q_r^m - 1 \right) + \alpha^m - \beta^m
\]

\[ = \frac{1}{\gamma} \left( \ln q_r^m - 1 + \frac{1}{q_r^m} \right) + \alpha^m - \beta^m
\]

\[ = \frac{1}{\gamma} \ln q_r^m + \alpha^m - \beta^m
\]

Then, when \( q_r^m > 0 \), the KKT conditions imply that

\[ \frac{1}{\gamma} \ln q_r^m + \alpha^m - \beta^m = 0 \Rightarrow q_r^m = \exp \left( -\gamma (\alpha^m - \beta^m) \right) \]

Consider another mode \( m' \) with \( q_r^m > 0 \), and the above condition also holds. Therefore, we derive the following equation:

\[ \frac{q_r^m}{q_r^{m'}} = \exp \left( -\gamma (\alpha^m - \beta^m) \right) \]

which implies that the passengers’ mode choice is in line with the multinomial logit distribution.

**Travellers’ destination choice follows multinomial logit distribution**

The procedure is similar to the preceding proofs, and we refer to Yang et al. (2000) for more details.

**References**


