Methodology for Transit Priority Lanes Design Problem Intended for Real Road Networks

Saeed Asadi Bagloee¹(∗), Majid Sarvi¹,
¹University of Melbourne, Australia;

Abstract

To promote public transport, we develop a practical method to identify best locations for transit priority lanes even in the heart of congested central business districts (CBDs). Inspired by the concept of Braess Paradox, we developed an index for each roads which indicates merit of a respective road to have a priority lane. This idea has been tailored to real life example. We use real data from the city of Melbourne which is one the largest transport network, as a case study. The same idea can also be extended to the freight mode aiming to enhance economic activities of the cities especially the CBDs.

1. Introduction

Traffic congestion is a chronic challenge in cities across the globe for which public transport is deemed to be an effective solution. It is for individuals rather than vehicles that the prioritization of mobility has to be determined (Mesbah et al., 2011; Zheng and Geroliminis, 2013). Accordingly, prioritization of the transit mode in terms of road space and signal time (Ahmed and Hawas, 2015; He et al., 2014; Hu et al., 2015; Ma et al., 2014) is of interest to practitioners as well as to academic theorists. In this study, we aim to address optimal reallocation of road space to transit mode on an existing urban transport network. In particular, we are interested in finding a network of exclusive transit priority lanes (in short, transit lane)¹ in the heart of cities in which prevailing traffic congestion makes the task more challenging and tedious. The concept of priority lanes has been introduced in many cities or for large scale crowd-gathering events (Black, 2004; Cova and Johnson, 2003; Smith and Hensher, 1998; Yingfeng and NaiQi, 2010; Zagorianakos, 2004) largely based on an engineering judgment call, not a systematic and solid decision making process. This study aims to fill such gaps.

A typical road can be modelled as a directional link consisting of one or more lanes. In general, roads and lanes can be used by all traffic modes (cars and buses). If a road is set to be a transit lane, it means that at least one of its lanes will be designated for the exclusive use of buses. As exceptions a transit lane can also be used by emergency vehicles and by high occupancy vehicles (HOV). Furthermore, priority lanes can also be designated for freight mode during off-peak hours so as to enhance urban economic viability. Freight mode has always been an underdog player in the transport system.

Expenses are invariably incurred from road marking, special signage and /or signals, lighting, etc. when assigning a lane to be a transit lane. Hence, the transit priority lanes design problem (TPLDP) can be articulated as a discrete network design problem (DNDP). When given a set of candidate roads, the problem emerges as to which roads should be selected for reassignment from their former purpose as a lane, while taking a limited budget into account. Diverting space from private modes in the interest of public transport is a delicate decision since it may adversely lead to more increased road congestion. Therefore in a bi-level design model in the primary (or upper) level one needs to minimize such adverse effects. One intuitive choice is to minimize the total travel time spent in the network, which is set to be the objective function. Due to such changes in the network, one needs to thoroughly capture the mutual interaction between transit mode and private modes and different vehicle classes of the private modes (cars, HOV, trucks etc). To this end, in the lower level, the model formulation should account for a multiclass and multimodal traffic assignment problem. Consequently the TPLDP (like a DNDP) is cast as a bi-level mixed integer, nonlinear optimization problem. It is proven that any bi-level programing problem is NP-hard so that the problem becomes quickly intractable as the size of the problem increases (Ben-Ayed and Blair, 1990). The aim is to develop a working methodology tailored to large sized networks (Bagloee and Ceder, 2011).

¹ Lanes can be designated for the exclusive use of buses and emergency vehicles (we collectively call them transit lanes). In fact the question of interest becomes: which roads should be selected to cease one lane to be used exclusively by the transit modes.
Therefore, real networks of Winnipeg, Canada and Melbourne, Australia which have over 150 and 2900 zones, respectively, are used as case studies.

2. Overview of the methodology

Though a variety of heuristic methods such as genetic algorithm have been introduced in the literature to tackle large sized networks, we develop a the branch-and-bound (BB) algorithm which aims to produce an exact solution. The steps of the methodology are specified hereafter. First, a set of roads with significant transit volume is identified as candidate roads. Second, using a BB method, the possibility of selecting a subset of the candidate roads is investigated, such that the overall performance of the traffic system including private and transit flows is not negatively impacted.

Incidentally, it is even possible, at this stage, to improve the overall performance of the network, for reasons well-illustrated by the Braess Paradox. The paradox shows how, as a result of closing road(s), overall traffic flows can improve. Empirical evidence, as well as mathematical theories have proven that the Braess Paradox is rampant in real transport networks (Bagloee et al., 2013; Braess et al., 2005; Nagurney, 2010; Roughgarden and Tardos, 2002). We then solve the TPLDP, while considering multimodal and multiclass features of the traffic flow which enhances the realism of the model. During an iterative loop process, the TPLDP is cast under a mode choice model to capture the elasticity of the travel demand amongst different modes and vehicle classes. In other words, according to the existing modal split the first TPLDP is solved. The outcomes (i.e. travel times of public and privates modes) are then fed back to the mode choice model to arrive at a new set of travel demands which are used to solve the next TPLDP. This process is conducted based on the method of Successive Averages (MSA) to ensure a guaranteed convergence over the demand matrices. For the modal split a nested logit formula is employed.

Given a set of candidates sorted based on their merit indices a pseudo algorithm of the methodology is provided as following:

Step 1, initialization: set iteration \(i=1\) and the existing travel demand as initial demand matrices denoted by \(D_i\). Also set a trivial value for \(\varepsilon >0\).

Step 2, Solving the TLPDP

Step 2-1, Initialize the BB’s tree by the first node representing the do-nothing scenario. Given the (updated) demand matrices \(D_i\), carry out traffic assignment for the existing network scenario (a scenario without any transit lane). Save the total travel time (TTT) as the incumbent value or the best upper bound value so far (UB*).

Step 2-2 based on the merit index and BB’s node selection rule find a node to do the branching. Create two new branches and nodes one representing for adding a candidate as priority lane \((y=1)\) and the other one for \(Y=0\). If there is no room for branching at this selected node (i.e. end of the tree at this node), carry out traffic assignment and update the upper bound UB*. Mark the solution corresponding to UB* as the best solution and terminate the TPLD algorithm (go to Step 3).

Step 2-3 For the selected node compute a lower bound denoted by LB. If \(LB>UB*\) then no better solution can be found so fathom (close down) the node and return to Step 2-2. Otherwise \((LB<=UB*)\) based on the merit index and BB’s branching rules, select a candidate to branch out two new nodes. Return to Step 2-2.

Step 3, run the modal split model to set up the demand matrices and denote the matrices with \(M_i\). If \(|(D_i-M_i)/M_i|<\varepsilon\), terminate the algorithm and return the marked solution (corresponding to UB*) as the final solution. Otherwise, do \(i=i+1\) and \(D_i=D_i+((1/i)\times(M_i-1-D_i))\), return to Step 2.

\(^2\) The heuristic methods struggle with a number of issues pertaining to the uncertainty of the degree of success of the results. They may achieve some good solutions, yet only optimal locally. The problem is that the success or lack of success of proposed solutions is effectively unforeseeable.

\(^3\) It is all about promoting public modes, such that some modal shift from private towards public mode is not surprising.

\(^4\) The do-nothing scenario (existing network) is a feasible solution to the TLPDP and it is possible that the problem might not have any better solution. It is important to note that, we force the problem not to return any priority lane at the cost of deteriorating the current level of
3. Innovative components of the methodology

In order to tackle large-sized networks, a number of initiatives have been developed: (i) for the candidate roads, upon the existing congestion levels and transit ridership concept of merit index is defined which is then employed in the BB algorithm to seek solutions, (ii) a particular search protocol is developed over the BB tree structure, such that the search for solutions becomes memoryless which saves a significant portion of the RAM, (iii) for the lower level problem of traffic assignment, we employ the SOLA\(^5\) (Second Order Linear Approximation) method which has recently been introduced to the industry as part of the EMME 4 (a leading transport planning software). In terms of the convergence speed of the multiclass traffic assignment, SOLA has shown significant superiority over state-of-the-art, bush-based algorithms (Florian and Morosan, 2014a) such as OBA (Bar-Gera, 2002) TAPAS (Bar-Gera, 2010) and algorithm-B (Dial, 2006) and LUCA (Gentile, 2009; Gentile, 2014). It is important to note that while multiclass traffic assignment solutions traditionally suffer from failing to yield unique solutions, recent research (Florian and Morosan, 2014b) shows that a sufficiently well-converged SOLA traffic assignment exhibits path flow and class link flow uniqueness. Furthermore, (iv) given the unprecedented sizes of the case studies, we also resort to parallel computation techniques for solving the multiclass traffic assignment problems. SOLA can naturally offer such important computation options (v) in traffic analysis, due to computational complexities, physical capacities of the roads are largely ignored. In order to enhance the realism of the modelling, we explicitly take roads’ capacities into account, applying a recently introduced method known as “Inflated Travel Time” (Bagloee and Sarvi, 2015). The capacity of the transit system (in terms of the number of buses and the number of seats available) is also explicitly considered.

4. Contribution to the literature

This paper contributes to the literature on a number of fronts. (i) For the first time, a network-wide approach to the TPLDP is tailored for real-sized networks. Traffic congestion as well as two important features of real traffic flow (multiclass and multimodal) are explicitly taken into consideration. (ii) The Braess Paradox is utilized to help nullify the adverse effect of transit priority lanes on the private mode by providing faster public transport. (iii) An especially RAM-efficient BB algorithm is proposed and is coded in-house, i.e., its simple structure can be easily embedded in any programming language. (iv) An efficient traffic assignment method (SOLA) is employed while explicitly accounting for road capacities. (v) In addition, the transit system is explicitly subjected to the fleet size constraints in order to arrive at a more realistic assessment. (vi) Due to anticipated impacts of the transit lane, the elasticity of travel demand in terms of modal shifts is also explicitly captured using logit formula.

5. Mathematical formulation of the transit priority lanes design problem (TPLD)

For ease of formulation, we adopt the following convention: a road denoted as a candidate by \(c\) (conceivably with three lanes), is replaced with two new roads – (i) road \(c \in \mathcal{A}\) with only one lane to be designated either as a mixed mode road or an exclusive transit lane or road, and (ii) road \(c \in \mathcal{A}\) with two lanes designated for mixed mode use. As such, transit lanes are separately denoted by one-lane roads\(^6\). Therefore, we can have: \(\mathcal{A}\) : set of roads currently with mixed modes (transit and private modes), but considered as a candidate for exclusive use by transit modes, and the rest of the roads are denoted by \(A\). Note, that in order to preserve the connectivity of the network, where speculations forfeit space to buses, the roads must have at least two lanes. Thus, should they be nominated to appropriate one lane as a transit lane, there will still be at least one remaining lane (hereafter, “candidate” will suffice for reference to a candidate road).

\(N\) : set of nodes.
\(B\) : budget available to cover the costs of transit lane implementations such as marking, pavement, curb raising, etc.


\(^6\) By doing so, we can alternatively call it a transit priority lane or transit road.
$y_a$: binary decision variable associated with candidate $a \in \overline{A}$; 1: to be used as exclusive transit lane and 0: to remain mixed use road or lane.

c_{a}$: implementation cost associated with candidate $a \in \overline{A}$.

$\pi_a$: public passenger volume on road $a \in \overline{A}$.

$x_a, \overline{x}_a$: private and transit traffic flow in passenger car equivalent or unit ("pcue") on road $a \in \overline{A}$ respectively (Note, the network available to the private and transit are $A$ and $A \cup \overline{A}$ respectively, hence $x_a, \overline{x}_a \geq 0$ for $a \in A$ and $x_a = 0, \overline{x}_a = 0$ for $a \in \overline{A}$). Note that $\overline{x}_a$ is passenger traffic volume on the road while $x_a$ is the car equivalent value of the corresponding number of buses on the respective road.

$t_a(y_a + \overline{x}_a)$: general travel time of link $a \in \overline{A}$, a non-decreasing BPR function of road’s flow $x_a + \overline{x}_a$ (called delay function (Sheffi, 1985; Spiess, 1990)). Note that $\overline{x}_a$ is the fixed value as background traffic.

$A_a, A_a^*$: set of links starting and ending at node $n$ respectively; $A_a, A_a^* \subseteq A \cup \overline{A}$.

$M$: set of distinct user classes.

$m_{ab}$: additional delay (constant bias) perceived by auto class $M_m \in M$. 

$m_{ax}$: traffic volume of auto class $M_m \in M$ of link $A_a \in \overline{A}$, in other words: $\sum_{a \in A_a} m_{ax} = \min(\sum_{a \in A_a} x_a, \overline{x}_a)$ (see Equation (6)) where $\overline{x}_a$ is the physical capacity of the respective road. Likewise, $\overline{x}_a$ is the capacity of the transit mode.

$R$: set of origin-destination (OD) pairs $R \subseteq N^2$.

$q^m_r$: travel demand in pcu for OD $r$ pertaining to the auto class $M_m \in M$.

$g_{ij}$: (public) passenger demand from node $i$ to destination node $j$. In order to simplify the notation, we define $g_{i} = \sum_{j \in \{N \setminus \{i\}} g_{ij}$, see (Spiess and Florian, 1989).

$P^m_r$: set of paths between OD $r$ available to the auto class $M_m \in M$.

$h^m_r$: traffic flow on paths $k \in P^m_r$, pertaining to the auto class $M_m \in M$.

$\delta^m_{a,K}$: road-path incident index, 1 if link $a \in A \cup \overline{A}$ belongs to path $k \in P^m_r$, pertaining to the auto class $M_m \in M$.

$w_n$: average waiting time at node $n \in N$ pertaining to transit system.

$f_a$: sum of frequency of service for all transit lines on link $a \in A \cup \overline{A}$.

The bi-level BLDP may be written as (all variables and parameters are considered non-negative unless otherwise stated):

$$
\begin{align}
\min \quad & \sum_{a \in \overline{A}} (x_a + \overline{x}_a) t_a(x_a + \overline{x}_a) \\
\text{s.t.} \quad & y_a = 1 \text{ or } 0, \quad a \in \overline{A} \\
& \sum_{a \in \overline{A}} c_{a} y_{a} \leq B \\
& \min \quad \sum_{a \in \overline{A}} \int_{0}^{t_a(x_a + \overline{x}_a)} dx + \sum_{m \in M} m_{ax} h^m \\
& \text{s.t.} \quad \sum_{m \in M} h^m = q^m, \quad r \in R \\
& x_a = \sum_{r \in R} \sum_{k \in P^m_r} h^m \delta^m_{a,K}, \quad a \in A \cup \overline{A} \quad \text{and} \quad x_a = \min(\sum_{m \in M} x_a^m, \overline{x}_a), \quad a \in A \cup \overline{A} \\
& \overline{x}_a \leq U, \quad a \in \overline{A} \\
& x_a \leq (1 - y_a) U, \quad a \in \overline{A} \\
& \overline{x}_a = \arg \min \sum_{m \in M} \sum_{a \in \overline{A}} \overline{x}_a f_a x_a + \sum_{n \in N} w_n, \\
& \text{subject to} \quad \sum_{a \in \overline{A}} \overline{x}_a - \sum_{a \in A} \overline{x}_a = g_{i}, \quad i \in N \\
& \overline{x}_a \leq f_a w_n, \quad \overline{x}_a \quad a \in A \cup \overline{A} \\
& \overline{x}_a \geq 0, \quad a \in A \cup \overline{A} \\
\end{align}
$$

As mentioned before the travel demand matrices are iteratively subjected to modal split model, for the sake of brevity we omitted an index representing the modal split iteration.
Equation (1) describes the upper level goal of minimizing the total travel time. Mathematical expressions (2) and (3) ensure the feasibility of the projects with respect to costs and available budget. At the lower level, mathematical expressions (4), (5), (6) the Beckmann formulation of UE flow pertaining to the private mode is computed. Constraints (2), (7) and (8) ensure that private flow will not enter the dedicated transit lanes \( U \) is a sufficiently large value, total demand \( \sum_{m} (q_{m}^x) \). Although Equation (7) is redundant, it is placed in the constraints to emphasize that buses can use candidate roads either exclusively (if it turns out to be \( y_{a} = 1 \)) or mixed with motorized mode (i.e. \( y_{a} = 0 \)). If it is decided that candidate \( a \) be an exclusive transit lane/road (i.e., \( y_{a} = 1 \)), then equation (8) ensures the respective road be closed to private mode (i.e. \( 1 - y_{a} = 0 \)). Equation (9) carries out capacity-constraint8 transit assignment based on optimal strategy (Cepeda et al., 2006; Spiess, 1993; Spiess and Florian, 1989) and it returns \( \tau_{x} \) as passenger traffic volume. The equation also returns effective frequency of the transit lines and alternatively number of buses on the roads. The equivalent value of buses in pcu (denoted by \( \pi_{x} \)) is then considered as background traffic in the traffic assignment (Spiess, 1984).

The multiclass facet of the traffic assignment is embedded in the interpretation of the bias term \( b_{x}^{*} \) in which all distinct auto classes using link \( a \) are subject to the same congestion level (based on the total traffic volume of all classes), plus an additional term, the bias term, exclusive to each class (i.e. \( b^{*}_{x} = \tau(x_{a}) + b_{a}^{*} \)). Such formulations define a simplified way to consider the multiclass aspect of traffic flow10.

The algorithm is written using Visual Basic linked with Ms-Excel as an interface and MS-Access to efficiently handle the data. It is also synchronized with EMME 4 to carry out the traffic assignment.

6. Numerical results

A prerequisite of transit planning network and design is a model to investigate different scenarios. To this end, the authors have developed a comprehensive, multimodal and multiclass transport model which is also computational efficient. The transport model consists of 2959 traffic analysis zones or zones, 26530 intersections or nodes as well as 67842 directional links or roads. It is a multimodal model (a model with private cars as well as public transport). The public transport network or transit network consists of 17 transit vehicle types, 817 transit lines and 47595 transit line segments. It is also a multiclass model that is the private mode consists of two matrices, the passenger car mode as well as freight mode or trucks.

Figure 1 show the extent of the model that cover a large swath of land including suburb and exurb. We run the algorithm and calculate merit index for every road of the network which is shown in Figures 2 and 3, covering the CBD and central area of the city of Melbourne. With the same token, we can compute merit index for the freight mode aiming to provide greater accessibility to heavy vehicles during off-peak hours. This has also been shown in Figure 4.

Note Eq. (9) is solved using method proposed by Cepeda et al (2006) in which the capacities of the buses are explicitly considered.

It is worth noting that the Beckmann formulation is convex w.r.t to \( x_{a} \). For each class, the shortest path computations of each class take into account the class-specific bias as well as the travel time given by the volume delay. Therefore, it is not necessary to store the class-specific volumes explicitly \( (s_{a}^{x}) \) - the total volumes are sufficient \( (x_{a}) \). Spiess, H. (1984) Contributions à la théorie et aux outils de planification des réseaux de transport urbain. Montréal: Université de Montréal, Centre de recherche sur les transports.

The way that a multimodal traffic assignment is conducted in EMME 4 is as follows: based on the headway and transit demand, prior estimation of transit volume is made (i.e. \( \pi_{x} \)). This \( \pi_{x} \) is then treated as background volume for the auto traffic assignment, followed by conducting a transit assignment to get a more precise assignment result. Accordingly, \( \pi_{x} \) in Eq (4) is treated as a constant term derived from transit assignment (Eq(9)). Since our primary intention was to make use of commercial software for the traffic assignment (i.e., EMME 4), we referred interested readers to the software manual, INRO (2015) EMME 4 v 4.1. EMME3 User’s Guide 4.1 ed, Montreal, Quebec, Canada. and Boyce, D. (2014) Network equilibrium models for urban transport. Handbook of Regional Science eds Fischer, M.M., Nijkamp, P. Springer Berlin Heidelberg, pp. 759-786.
Figure 1 the EMME 4 model developed for the city of Melbourne

Figure 2 Melbourne, the CBD, public transport, transit priority lane based on the merit index shown as thickness
7. Conclusions and outlook for full paper

The full paper provide a more detailed analysis, both from a theoretical and experimental viewpoint, including numerical results of the colossal network of Melbourne (+2900 zones). More specifically, we establish an intuitive way to target Braess tainted roads to be used as candidate roads, and to be used in the BB algorithm. Likewise, we extend the idea of transit lane to the HOV and freight modes. The proposed methodology in terms of the theoretical features and real applications may be of interest to scholars in transportation and optimization fields as well as practitioners and traffic authorities in the industry.
References


