Optimal allocation of designated bus lanes in multi-modal urban networks

Dimitrios Tsitsokas

Mohammadreza Saeedmanesh

Anastasios Kouvelas

Nikolas Geroliminis

École Polytechnique Fédérale de Lausanne **May 2018**

STRC | 18th Swiss Transport Research Conference
Monte Verità / Ascona, May 16 – 18, 2018

École Polytechnique Fédérale de Lausanne

Optimal allocation of designated bus lanes in multi-modal urban networks

Dimitrios Tsitsokas, Mohammadreza Saeedmanesh, Anastasios Kouvelas, Nikolas Geroliminis
Urban Transport Systems Laboratory
École Polytechnique Fédérale de Lausanne
GC C2 390, Station 18, 1015 Lausanne, Switzerland
phone: +41-21-69-32481
fax: +41-21-69-32431
{dimitrios.tsitsokas, mohammadreza.saeedmanesh, tasos.kouvelas, nikolas.geroliminis}@epfl.ch

May 2018

Abstract

Road space distribution among multiple modes of transport in contemporary urban environments has attracted the interest of several researchers and policy planners. Revealing the relation between road space share and global performance of the network would result in more efficient transport system design. The introduction of dedicated bus lanes in parts of the network has been proposed as a measure to allow high occupancy vehicles to travel through regions with high traffic load without long delays. In this way, the bus-system operates more efficiently and is more competitive to the private car option in terms of travel time. This effect can stimulate a significant mode shift to bus use and alleviate congestion by reducing the number of low occupancy vehicles. In this work we address the problem of optimal allocation of exclusive bus lanes in a given multi-region multi-modal urban network with the aim of minimizing the total passenger hours travelled (PHT). A queueing theory based traffic flow model, known as the Store-and-Forward (SF) model, is utilized to simulate the evolution of congestion inside the network by keeping track of the queues inside all links over time. Our goal is formulate this problem in a Mixed-Integer Linear Programming (MILP) form and attempt to solve it to optimality.

Keywords

Exclusive Bus Lanes, Store-and-Forward, Transit Priority

1 Introduction

Traffic congestion in densely populated cities is a problem of increasing significance, which hinders mobility and delays commuters in many metropolitan areas around the world. Resulting from rapid urbanization and private vehicle ownership rise, the concentration of vehicles in central roads and arterials during peak-hours very often exceeds the network capacity and leads to gridlocks that significantly prolong travel time. Building new infrastructure such as new roads or rail systems to accommodate the excessive demand is usually not a viable option due to multiple topological, financial and political constraints. In such cases, traffic planners focus on improving the efficiency of the existing public transit system through a variety of different measures. Their goal is often to motivate commuters to swift from private vehicle to public transit use during peak-hours, thus alleviating the congestion problem by reducing the number of circulating vehicles.

There is a growing interest in the enhancement of public transit systems in many cities around the world facing traffic congestion. A review of the developments on bus rapid transit (BRT) systems implemented in several cities around the world can be found in Deng and Nelson (2011). The introduction of exclusive bus lanes (EBLs), a key asset of most BRT systems, has been seen as a cardinal bus-priority measure in many cities around the world. The idea behind this is that since buses carry a larger number of passengers compared to private cars, they should be allowed dedicated space to move through traffic congestion without delay. This policy is believed to boost a significant improvement to the bus system in terms of travel time and reliability, which can cause passengers to swift from car to bus use during peak hours.

The space allocation problem that rises in the design of an EBL network can be very challenging. Introducing dedicated bus lanes to a road may ensure a smooth flow of buses through congestion but in the same time, this measure reduces the capacity of the road for the rest of the traffic and thus increases the probability of a gridlock. In other words, if an extensive length of road space is given to EBLs, the network may reach a congested state more often and for longer periods than in the case of no EBL setting. Underused road space and increased traffic congestion due to the existence of high-occupancy vehicle (HOV) lanes in California has been reported in Chen *et al.* (2005). Such an effect is more likely to appear if the expected mode swift from private cars to public transit is negligible. On the other hand, too little space given to EBLs would hardly improve the public transit efficiency, as buses would still be slowed down due to traffic congestion in most of their routes. Hence, it is crucial to understand how the number and the spatial distribution of exclusive bus lanes inside a traffic network influence to the overall performance of a multi-modal traffic network.

This work tackles the problem of identifying the optimal distribution of exclusive bus lanes in a given urban traffic network where a fixed bus system operates. Our objective is to decide whether an EBL should be installed on every road of the network with at least one bus line or not. The first step to reach this objective is to identify a reliable model to assess the performance of the network in a microscopic level for any feasible candidate solution (EBL layout). The Store-and-Forward model, which is based on queueing theory principles, is used to simulate the dynamics of traffic congestion inside the network and plays the role of our simulator in this work. The second step is to formulate an optimization framework incorporating this model in order to find the optimal EBL layout, i.e. the set of links with an exclusive bus lane, which lead to maximum network performance in terms of travel time for passengers. Our goal is to formulate this problem in a Mixed Integer Linear Programming form and try to solve it to optimality for a case study of the San-Francisco network.

2 Background

A considerable amount of research works on road space distribution and transit priority schemes in urban multi-modal networks can be found in the literature. For instance, Waterson *et al.* (2003) evaluated the effect of imposing strong bus priority measures on the behavioral responses of commuters recognizing the potential harm they can cause on the public transit system.Dahlgren (1998) evaluated the impact of constructing a High Occupancy Vehicle (HOV) lane aiming at identifying the conditions under which a HOV lane would be beneficial to overall traffic performance in a freeway. Daganzo and Cassidy (2008) analyzed how freeways are used by competing transport modes and highlighted the necessity of considering the occupancy of each mode as a weighting factor in the decision making process regarding road space allocation. Gonzales *et al.* (2010) analyzed the interaction of different transportation modes in an aggregated way by utilizing the MFD modeling approach and showed that dedicating space to high occupancy modes of transport can improve the global performance of the network. More recently Gonzales and Daganzo (2012) analyzed system optimum solutions for space distribution in multi-modal networks for the morning commute problem. They also identified the car capacity reduction of the network resulting from the dedicated space for public transit, thus highlighting the need to balance the trade-off between transit enhancement and car traffic disturbance. In their works, Farid *et al.* (2015) and Farid *et al.* (2018) studied several transit preferential treatment strategies as a way to improve transit system in a person-based level, by focusing on the evaluation of exclusive bus lanes and queue jumper lanes via analytical and simulation models.

Several researchers have formulated the EBL setting problem in the form of a bi-level programming model. Mesbah *et al.* (2011) formulated transit priority problem in a bi-level optimization-programming model, which operates similar to a leader-follower Stackelberg game. The upper level includes the objective function from the system managerś perspective while the lower level includes the mode choice and the traffic and transit assignment from the userś perspective. Similarly, Bingfeng *et al.* (2017) proposed a bi-level programming model where the upper level defines the EBL layout which determines the link impedance functions, while the lower level solves a user equilibrium assignment problem. Branch-and-bound algorithm is used to solve the optimization problem with the aim of minimizing total passenger travel cost. Sun and Wu (2017) and Yu *et al.* (2015) propose similar bi-level models and include additional objectives related to pollutants emissions, transit operating cost and passenger waiting time in bus stops. All the above models consider static traffic conditions and simulate travel time by making use of the Bureau of Public Roads (BPR) or other similar types of functions.

Zheng and Geroliminis (2013) proposed a novel macroscopic approach for determining the optimal road space distribution among competing modes of transport in a multi-region network based on the Macroscopic Fundamental Diagram (MFD) of the regions. This modeling captures the dynamics of congestion in the region level and describes the objective function in terms of passenger total travel time. The highly non-linear optimization problem is solved by using a Sequential Quadratic Programming method. The outcome quantifies the fraction of the regional road space that should be assigned to exclusive bus lanes to achieve minimum passenger travel time. The present work is stimulated from the above research and attempts to solve the same optimization problem but in a microscopic level, by investigating the impact of the spatial distribution of the exclusive bus lanes on the total network performance.

It is clear that the reliability of a proposed solution for the problem of exclusive bus lanes distribution directly depends on the accuracy and robustness of the model used to simulate the evolution of the traffic flow inside the network. Many of the existing studies for the EBL problem step on the concept of BPR functions to compute the travel time inside the links of the network. However, this concept assumes static conditions and cannot capture the dynamic nature of congestion propagation and the spillback effects. This is why queueing theory based models are considered to be more realistic in simulating network traffic flow. The Store-and-Forward method, which belongs to this category, widely used today mainly in telecommunication networks, has been utilized in research studies on network control strategies in Aboudolas *et al.* (2009), Aboudolas *et al.* (2010), as also in Varaiya (2013) and Kouvelas *et al.* (2014). A more enhanced modelling technique based on the same concept is known as the S-model and it has been employed in several works, as in Lin *et al.* (2010). A similar logic queueing modeling with finite capacity of queues, that is able to analyze the sources and effects of congestion, is proposed by Osorio and Bierlaire (2009) and applied in a real hospital service queueing problem.

In this work, in an attempt to take into account the dynamics of the congestion propagation within the network, we utilize the Store-and-Forward (SF) model to keep track of the length of the queues of vehicles that are present inside every link in all time steps of the simulation test. This is a way to assess the performance of the network in terms of the total time that travelers spend inside the network by integrating the queues over time. Road capacities are considered as an input information for SF and are defined by the proposed EBL distribution plan, i.e. the roads on which we consider the existence of an exclusive bus lane display decreased capacities to accommodate regular traffic. An optimization framework enclosing the SF simulation model is formulated to locate the best EBL plan in terms of total passenger hours travelled (PHT) under predefined dynamic traffic demand. A Mixed Integer Linear Programming (MILP) formulation of the problem is also a goal of the present research endeavor.

3 Methodology

3.1 Problem Description

Consider an urban traffic network represented as a directed graph which consists of a set of links (roads) L and a set of nodes (junctions) J. Every link is identified by its two edges (starting node and ending node). The length and the number of lanes for each link are considered known. In addition, every junction has a fixed and known traffic signal plan. There are two modes of transport using this network: buses and private cars. The operational characteristics of the installed bus system (routes and frequencies) as also the average passenger occupancy of each line are fixed in the context of the present work. Any link of the network can be an origin and/or a destination of a trip. An O-D time-dependent demand matrix feeds the network with private car flow. Our goal is to decide which links should have one dedicated bus lane in order to achieve minimum total travel time for all passengers during a specified time horizon representing the morning/evening peak.

In order to make decisions about the setting of an EBL in each of the links, we need a way to assess the network performance for every candidate EBL distribution in terms of the total time that travelers spend inside the network for a defined time horizon. The main impact of the introduction of an exclusive bus lane on a link is that its capacity gets decreased which means that the link can serve lighter car flow before it gets congested comparing to the previous state of no bus lane existence. The capacity of a link is a function of the link's number of lanes and length. Therefore, the setting of an exclusive bus lane decreases by one the number of lanes that are available for the rest of the traffic. This can be modelled by the introduction of a binary variable Y which is equal to 1 if a bus lane exists in this road or 0 otherwise. Thus, for a link z, we can compute the capacity as $C_z = (n_z Y_z)l_z/l_{v,avg}$, where n_z and l_z is the number of lanes and length of link z respectively, Y_z is the binary variable indicating the existence or absence of an EBL, and $l_{v,ang}$ is the considered average vehicle length. If a link becomes congested spillbacks are propagated backwards in a way that depends on the network topology, the respective accumulations and the capacities of the neighboring links. To take into account the dynamics of congestion propagation, the queueing theory based Store-and-Forward (SF) model is utilized to simulate the evolution of traffic queues in every link of the network subject to a time-dependent traffic demand. The details of this model are discussed in the following section.

3.2 The Store-and-Forward model

The Store-and-Forward model simulates the traffic flow inside the network by calculating the number of vehicles (queues) inside the links based on a conservation equation. The road network is represented as a directed graph, which consists of a set of links Z and a set of nodes J. Links are connected to other links through nodes. For each node (or junction) $j \in J$, we define a set of incoming and outcoming links, I_j and O_j respectively. For every junction j, the cycle time C_j and the green times $g_{i,o}$ (for the flow going from incoming link i to outcoming link o) are constant throughout this work. Consider a link z connecting two junctions $M, N \in J$ such that $z \in O_M$ and $z \in I_N$. This means that the vehicles move from node M to node N through link z. The dynamics of link z are described by the following conservation equation:

$$x_z(k+1) = x_z(k) + T[q_z(k) - s_z(k) + d_z(K) - u_z(k)] \tag{1}$$

where $x_z(k)$ is the number of vehicles being present inside the link z at time kT, T is the time step of the simulation test and k is the time step index. The terms q_z and u_z represent the total inflow and outflow of link z respectively and the terms s_z and d_z represent the exit flow and the generated demand flow of link z respectively. The exit flow is calculated as $s_z(k) = t_{z,0}q_z(k)$, where $t_{z,0}$ is the exit rate of link z which is considered as known. The generated demand flow $d_z(k)$ of the link z for every time step k is also known by the dynamic O-D demand matrix, which is an input for this problem. The incoming flow is calculated as $q_z(k) = \sum_{w \in I_M} u_{w,z}(k)$, where $u_{w,z}$ is the outflow of the link w that moved to link z at time step k. If a move from an inflow link to an outflow link is not allowed, this term will be equal to zero. This will be clarified below. The queue in every link z cannot be greater than the maximum queue that the link can host based on its available surface. This is described by the following constraints:

$$0 \leq x_z(k) \leq x_{z,max}, \forall z \in Z \tag{2}$$

where $x_{z,max}$ is the maximum acceptable queue of link z. This value is equal to the capacity C_z of the link, calculated as $C_z = (n_z l_z)/l_{v,avg}$, where n_z and l_z are the number of lanes and the length of link z respectively and $l_{v,av}$ is the average vehicle length. The constraints of Eq. (2) protect the downstream links from oversaturation during periods of high demand by limiting the inflow from upstream links in cases of congestion. A core simplification of this modelling approach, as introduced in the paper of Aboudolas *et al.* (2009), is the following: assuming that there is enough available space in the downstream link w, as controlled by the Eq. (2), the fraction of the outflow of link z that is transferred to the link w, noted as $u_{z,w}$, is equal to the saturation flow S_z of link z for as long as link z has right of way towards link w and zero otherwise. To remove the complexity imposed by the traffic signal phase, we can say that traffic flows continuously from z to w in an average rate equal to the saturation flow computed for the connection of the two links. In this case, the outflow from z to w is described as:

$$u_{z,w}(k) = \frac{S_{z,w} g_{z,w}}{C_j}, \tag{3}$$

where $z \in I_j$ and $w \in O_j$. In this work we calculate the value of the saturation flow for every pair of connecting links z and w of a junction j based on their number of lanes and on the turning rates that apply for each particular connection. By definition, the saturation flow of a road is equal to a maximum passing rate in free flow conditions (usually taken equal to 1800 veh/h per lane) multiplied by the number of lanes of the road. However, in many cases, not all lanes of a road lead to a certain direction (e.g. left turn) while it is possible that the downstream link may have fewer lanes than the upstream. Also, in the case of a road with lanes allowing vehicles to move to several directions in a junction (e.g. straight and right turn), the saturation flow for a specific connection (e.g. from z to w) is also dependent on the fraction of the total traffic of the upstream link wishing to travel on this connection (defined by the turning ratio $t_{z,w}$). Based on these thoughts, we calculate the saturation flow $S_{z,w}$ as the minimum of three quantities:

$$S_{z,w} = min\{n_{z,w}, n_w, n_z t_{z,w}\} * 1800 \text{ (veh/h)} \tag{4}$$

where $n_{z,w}$ is the number of lanes of link z that allow vehicles to go to link w, n_w is the number of lanes of the downstream link w, n_z is the total number of lanes of the upstream link z and $t_{z,w}$ is the turning rate from link z to w. In other words, the first term represents the saturation flow that can exit link z and head to link w, the second term represents the saturation flow that link w can accept and the third term represents the fraction of the saturation flow of the total link z with direction to link w.

Combining Eq. (2) and Eq. (3), the outflow $u_z(k)$ can be calculated from the following equations (disregarding the constraints Eq. (2)):

$$u_z(k) = \sum_{w \in O_N} u_{z,w}(k) \tag{5}$$

$$u_{z,w}(k) = \begin{cases} 0 & \text{if } x_w(k) \geq x_{w,max} \\ min\left\{\frac{S_{z,w} g_{z,w}}{C_j}, \frac{x_z(k) t_{z,w} g_{z,w}}{C_j}\right\} & \text{else} \end{cases} \tag{6}$$

$$x_{w,max} = C_w = \frac{n_w * l_w}{l_{v,avg}} \tag{7}$$

Indeed, according to Eq. (6) the outflow $u_{z,w}(k)$ is zero if link w is full at time kT. Otherwise, it is equal to the fraction of the current sum of vehicles $x_z(k)$ inside link z that wish to travel to link w if it less than the maximum possible traffic that can travel from z to w, which is the saturation flow $S_{z,w}$ for this connection. Finally Eq. (5) calculates the total outflow of link z by summing the outflows calculated for every downstream link of the junction N. Note that if any link $w \in O_N$ do not allow incoming flow from link z, then $u_{z,w}$ will be equal to zero based on Eq. (6), as the respective turning rate $t_{p,z}$ will be zero.

Based on this model, it is possible to have a picture of the congestion propagation inside the network in the form of the length of the queues (accumulation of vehicles) that are calculated for every link at every time step of the simulation test, as far as the turning rates for every pair of incoming-outcoming links and for every intersection, as also the exit rates in every link are known. In this work, the model will be applied to a case study where these rates are taken as an input from another traffic simulating software. The significance of using the SF modeling lies at its ability to evaluate the performance of the network under dynamic traffic conditions based on the dynamic evolution of the queues, and thus, evaluate different candidate EBL layouts at a low computational cost.

3.3 Optimization Framework

The Store-and-Forward model is utilized to simulate traffic flow inside the network over a predefined time. The network characteristics and the dynamic demand pattern are the inputs for SF model. The output of a simulation run is the evolution of queues for every link at every time step. This metric can be used to calculate the total passenger hours travelled (PHT) for the simulation time horizon for the considered demand and network characteristics. The queues provided by the SF model refer to the mixed traffic. Travel time for buses on dedicated bus lanes can be computed separately, by assuming that on an EBL buses move by a free flow speed.

The objective of this optimization problem is to decide which of the links $z \in Z$ should have an exclusive bus lane in order to maximize the network performance in terms of travel time for commuters. As described above, this is modelled by a binary variable Y which is equal to 1 if a bus lane exists in this road or 0 otherwise. This means that the decision variables are a set of binary variables $Y = \{Y_z \mid z \in Z\}$. If specific conditions must apply for a road in order to have an EBL, a subset Q of Z of all eligible roads is defined and our decision variables are $Y = \{Y_z \mid z \in Q \subseteq Z\}$. In the context of the present work, it is assumed that every link can either have no exclusive bus lane or only have one, which will be the most right lane on the direction of traffic. Two or more EBLs cannot be placed on the same link. Setting one exclusive bus lane on a road of the network translates in a decrease of the road capacity, as indicated by Eq. (7), since the number of lanes l_w which are available for the regular traffic is decreased by one.

4 Case Study

In order to test the developed methodology we perform simulation experiments for a real urban network with many bus lines. In this work, the Downtown area of San Francisco, in California is used as a case study. This network contains multiple bus lines that run in high frequencies during the peak hour, allowing for carrying high passenger loads that can assist in alleviating the traffic congestion in the city. A realistic simulation set-up of this multi-modal network is described in the current section.

4.1 Network description

The test site is a 6.5 square kilometer area of Downtown San Francisco (Financial District and South of Market Area), including 156 intersections and 426 links with lengths varying from 70 to 400 meters (see Figure 2(a) for the map representation of the area). The number of lanes per link varies from 1 to 6 and the free flow speed is 50 kilometers per hour. In the network there are 92 signalised intersections and 64 that do not have traffic lights. Traffic signals are all multiphase fixed-time, operating on a common cycle length of 90 seconds for the west boundary of the area (The Embarcadero), and 60 seconds for the rest of the network.

A microsimulation model exists for this area in the commercial software Aimsun (see Figure 2(b)). The model includes all the bus lines of the Downtown area of San Francisco, the timetables with the frequencies during the peak hour, as well as a realistic detailed OD demand table that has been calibrated with real data. In this area of San Francisco, there are 29 bus lines that run in regular frequencies (e.g. every 10 minutes) and traverse different number of links and bus stops. Out of the 426 urban links, 218 do not serve any bus line, and for the rest 208 the number of associated bus lines varies from 1 to 10 (see Figure 2(c)).

For the simulation tests, we assume that in the base scenario there are no designated bus lanes in the network. The research question is to decide the optimal location of the dedicated bus lanes, so as to maximize the total passenger hours travelled (PHT). The simulation step for the microscopic simulation model of the test site, is set to 0.5 seconds. Furthermore, in order to simulate the adaptivity of drivers and account for route choice effects in the OD demand, the Dynamic Traffic Assignment (DTA) module (C-Logit route choice model) is activated in Aimsun every 3 minutes, a time interval that is consistent with the average trip length in the test area of San Francisco.

4.2 Simulation set-up for store-and-forward model

From the microsimulation network we can extract all the necessary data that are needed as input for the store-and-forward model. First of all, we need static data for the topology of the network (i.e. number of lanes, coordinates, connectivity of links), as well as the signal settings per node and per movement (e.g. left, right, through). Moreover, in order to obtain dynamic data from Aimsun (e.g. turning ratios) we need to run some replications and record the data. In this case, we use time varying turning ratios (i.e. averages that change every 15 minutes). Another input for store-and-forward that needs to be extracted from Aimsun is the demand in the origins of the network.

5 Discussion and future steps

This paper addresses the problem of finding an optimal distribution of exclusive bus lanes in an urban multi-modal traffic network, which minimizes the total travel time for passengers. Our goal is to formulate this problem as a Mixed Integer Linear Program and try to solve it to optimality. The Store-and-Forward traffic flow model is utilized to simulate the dynamics of congestion propagation in the network. A microsimulation of San-Francisco network in the AIMSUN environment provides the Store-and-Forward model with the turning ratios, which determine the route choice of the travellers. Mode choice as also the mode shift from car to bus is not modelled at this point.

At this point we do not have any result as this work is in a preliminary stage. We are interested in investigating the characteristics of the solutions with the best values of the objective function, the spatial correlation of the selected links with bus lanes as also the results of a microsimulation test with AIMSUN for the best solution. Our goal is to formulate the problem as an MILP and try to solve it to optimality. It is possible that in this form the problem has too many variables and the computational cost becomes too large in the case of a real network. In this case, we can focus on methods to relax the problem in order to achieve better run time. Modelling dwell times of buses as also the mode choice will be our next priority.

6 References

Aboudolas, K., M. Papageorgiou and E. Kosmatopoulos (2009) Store-and-forward based methods for the signal control problem in large-scale congested urban road networks, *Transportation Research Part C: Emerging Technologies*, **17** (2) 163–174.

Aboudolas, K., M. Papageorgiou, A. Kouvelas and E. Kosmatopoulos (2010) A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road networks, *Transportation Research Part C: Emerging Technologies*, **18** (5) 680–694.

Bingfeng, S., Z. Ming, Y. Xiaobao and G. Ziyou (2017) Bi-level programming model for exclusive bus lanes configuration in multimodal traffic network, *Transportation research procedia*, **25**, 652–663.

Chen, C., P. Varaiya and J. Kwon (2005) An empirical assessment of traffic operations, paper presented at the *Proceedings of the 16th International Symposium on Transportation and Traffic Theory*, 105–124.

Daganzo, C. F. and M. J. Cassidy (2008) Effects of high occupancy vehicle lanes on freeway congestion, *Transportation Research Part B: Methodological*, **42** (10) 861–872.

Dahlgren, J. (1998) High occupancy vehicle lanes: Not always more effective than general purpose lanes, *Transportation Research Part A: Policy and Practice*, **32** (2) 99–114.

Deng, T. and J. D. Nelson (2011) Recent developments in bus rapid transit: a review of the literature, *Transport Reviews*, **31** (1) 69–96.

Farid, Y. Z., E. Christofa and J. Collura (2015) Dedicated bus and queue jumper lanes at signalized intersections with nearside bus stops: Person-based evaluation, *Transportation Research Record: Journal of the Transportation Research Board*, (2484) 182–192.

Farid, Y. Z., E. Christofa and J. Collura (2018) An analytical model to conduct a person-based evaluation of transit preferential treatments on signalized arterials, *Transportation Research Part C: Emerging Technologies*, **90**, 411–432.

Gonzales, E. J. and C. F. Daganzo (2012) Morning commute with competing modes and distributed demand: user equilibrium, system optimum, and pricing, *Transportation Research Part B: Methodological*, **46** (10) 1519–1534.

Gonzales, E. J., N. Geroliminis, M. J. Cassidy and C. F. Daganzo (2010) On the allocation of city space to multiple transport modes, *Transportation planning and technology*, **33** (8) 643–656.

Kouvelas, A., J. Lioris, S. Fayazi and P. Varaiya (2014) Maximum pressure controller for stabilizing queues in signalized arterial networks, *Transportation Research Record: Journal of the Transportation Research Board*, (2421) 133–141.

Lin, S., B. De Schutter, Y. Xi and H. Hellendoorn (2010) Model predictive control for urban traffic networks via milp, paper presented at the *American Control Conference (ACC), 2010*, 2272–2277.

Mesbah, M., M. Sarvi and G. Currie (2011) Optimization of transit priority in the transportation network using a genetic algorithm, *IEEE Transactions on Intelligent Transportation Systems*, **12** (3) 908–919.

Osorio, C. and M. Bierlaire (2009) An analytic finite capacity queueing network model capturing the propagation of congestion and blocking, *European Journal of Operational Research*, **196** (3) 996–1007.

Sun, X. and J. Wu (2017) Combinatorial optimization of bus lane infrastructure layout and bus operation management, *Advances in Mechanical Engineering*, **9** (9) 1687814017703341.

Varaiya, P. (2013) Max pressure control of a network of signalized intersections, *Transportation Research Part C: Emerging Technologies*, **36**, 177–195.

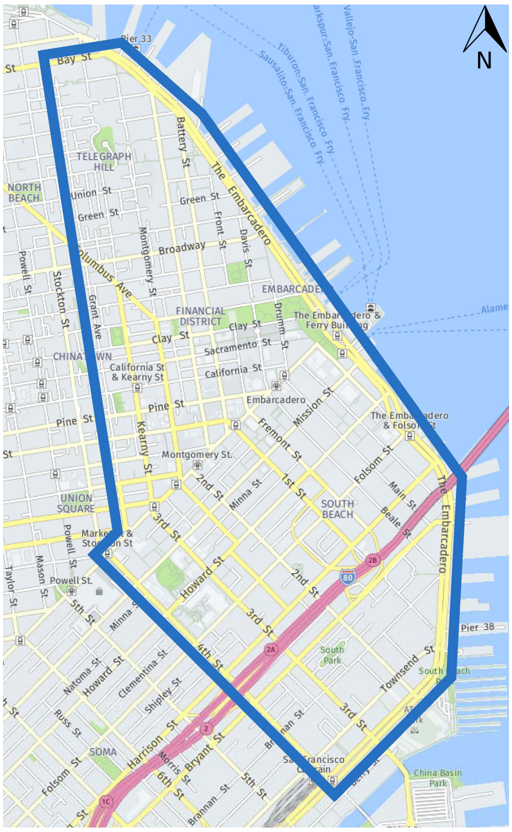
Waterson, B., B. Rajbhandari and N. Hounsell (2003) Simulating the impacts of strong bus priority measures, *Journal of Transportation Engineering*, **129** (6) 642–647.

Yu, B., L. Kong, Y. Sun, B. Yao and Z. Gao (2015) A bi-level programming for bus lane network design, *Transportation Research Part C: Emerging Technologies*, **55**, 310–327.

Zheng, N. and N. Geroliminis (2013) On the distribution of urban road space for multimodal congested networks, *Transportation Research Part B: Methodological*, **57**, 326–341.

Figure 1: San Francisco test network

(a) Map of San Francisco Downtown area



(b) Microsimulation model in Aimsun

(c) Number of bus lines traversing each link of the network