

Research Article

A Two-Layer Network Dynamic Congestion Pricing Based on Macroscopic Fundamental Diagram

Bangyang Wei ^{1,2} and Daniel(Jian) Sun ^{1,2}

¹State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai, China

²Center for UAV Applications and ITS Research, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, China

Correspondence should be addressed to Daniel(Jian) Sun; danielsun@sjtu.edu.cn

Received 22 February 2018; Revised 15 May 2018; Accepted 27 May 2018; Published 1 August 2018

Academic Editor: David Z. W. Wang

Copyright © 2018 Bangyang Wei and Daniel(Jian) Sun. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Dynamic congestion pricing has attracted increasing attentions during the recent years. Nevertheless, limited research has been conducted to address the dynamic tolling scheme at the network level, such as to cooperatively manage two alternative networks with heterogeneous properties, e.g., the two-layer network consisting of both expressway and arterial network in the urban areas. Recently, the macroscopic fundamental diagram (MFD) developed by both field experiments and simulation tests illustrates a unimodal low-scatter relationship between the mean flow and density network widely, providing the network traffic state is roughly homogeneous. It reveals traffic flow properties at an aggregated level and sheds light on dynamic traffic management of a large network. This paper proposes a bilevel programming toll model, incorporating MFD to solve the unbalanced flow distribution problem within the two-layer transportation networks. The upper level model aims at minimizing the total travel time, while the lower level focuses on the MFD-based traffic assignment, which extends the link-based traffic assignment to network wide level. Genetic algorithm (GA) and the method of successive average were adopted for solving the proposed model, on which an online experimental platform was established using VISSIM, MATLAB, and Visual Studio software packages. The results of numerical studies demonstrate that the total travel time is decreased by imposing the dynamic toll, while the total travel time savings significantly outweigh the toll paid. Consequently, the proposed dynamic toll scheme is believed to be effective from both traffic and economic points of view.

1. Introduction

Congestion pricing has been regarded as an effective traffic demand management policy that has been applied in many cities, among which Singapore was the first one to impose road congestion pricing at city wide level [1]. The development process has experienced two stages from the initial regional license scheme to the real-time dynamic electronic charging system. In United States, most tolled roads are related to new lanes or lanes that have been opened to High Occupancy Vehicles (HOV). In California, private investors have built the 91st fast lane based on the existing state highway S91, providing drivers the option of using an uncharged lane or a time-varying fee lane. These successful road congestion pricing schemes have aroused wide considerations in exerting

more efficient and impartial pricing strategies in large urban networks.

The theory of congestion charging was developed by Walters [2] and Vickrey [3] that the users should pay for both the toll equal to their own travel cost and the additional congestion cost imposed on other users due to the extra travel occupying the public road resource, which is consistent with the concept called “marginal cost” in economics. Sheffi [4] found that the additional cost added in the static models can drive the user equilibrium pattern to achieve system optimum. Ferrari [5] studied the charging problem for urban transport network with elastic demand and link capacity constraints and found that the model has no solution which satisfies the capacity constraints, but additional costs can be imposed on some links so that an equilibrium solution

that satisfies the capacity constraints always exists. In terms of multiuser classes, Leurent [6] designed a model that allows for differentiating travelers by means of an attribute called value of time that is continuously distributed over a real interval. Yang and Lam [7] and Yang and Bell [8] studied congestion pricing in the queue network under fixed and elastic demand, respectively, through establishing the bilevel programming model, where the lower level problem describes users' routes choice behavior under condition of both queuing and congestion, while the upper level problem is to determine road toll to optimize a desired system's performance. In view of the time-varying traffic demand, the static toll models need to be extended to dynamic toll models. Vickrey [9] applied deterministic queuing theory to firstly propose the dynamic pricing scheme called the bottleneck model that the travelers should pay the toll for eliminating the queuing delay without toll. Yang and Meng [10] combined the application of the space-time expanded network representing the time-varying traffic flow and the conventional network equilibrium modeling technologies. Carey and Watling [11] extended the formulation that used Cell Transmission Model (CTM) for constructing system optimum (SO) formulation to allow more general nonlinear flow-density functions and found that if the tolls computed from the dynamic system optimum (DSO) solution are imposed on the users, the DSO solution would also satisfy the criteria for a dynamic user equilibrium (DUE). Ban and Liu [12] formulated a link-nodes discrete-time dynamic toll model as a bilevel problem, where the upper level is to minimize the total weighted travel time, while the lower level is to capture the users' route choice behavior. Laval and Castrillón [13] analyzed the time-dependent congestion pricing scheme on two alternative routes that have different bottleneck capacities and surprisingly found that the minimum total system delay can be achieved with many different pricing strategies, which are valuable for traffic management. Although both the static and dynamic settings and cases that have been addressed so far are encouraging, there are some certain deficiencies: (i) traffic system is not in steady state at link level validated by simulation evaluation and field experiment, and thus toll calculated based on ideal curve is not optimal [14, 15]; (ii) existing studies on dynamic congestion pricing have been limited to certain idealized networks, because the model and computation of the link-based dynamic traffic assignment (DTA) are so complex that these research cannot be applied to a large city network practically.

Recent findings on the traffic dynamics at the network level named macroscopic fundamental diagram (MFD) have aroused wide attentions. Geroliminis and Daganzo [14, 15] used both detector data from microscopic simulation test for the San Francisco Business district in California and a data combination of fixed detectors and floating vehicle probes as sensors from field experiment in downtown Yokohama in Japan to find that MFD relating the number of vehicles in network to space mean flow or outflow presents a unimodal low-scatter relationship, if the road network is in homogenous congestion state. Scholars mainly focus on the research about its nature, influence factors, and application. Leclercq and Geroliminis [16] took advantage of the topology of the route

and the variation theory to obtain MFD, in order to overcome the homogeneity rule which points out the congestion state at link level should keep pace. The influencing factors on the shape of MFD have also been investigated gradually. Geroliminis and Sun [17] found the "hysteresis phenomenon" that a clockwise loop exists between the flow and density instead of linear curve and then provided the explanation with density inhomogeneity and lack of adequate data. In terms of the influence of turning flow in intersections on MFD, Geroliminis et al. [18] found that the left traffic flow reduces the largest value of MFD. Because of the increasing recognition about MFD, researchers have started to focus on how to apply MFD to region traffic management. One of the main tasks was to investigate network division, thus obtaining a well-defined MFD. The most classic method was developed by Ji and Geroliminis [19] that divided the entire network according to the congestion feature [20, 21], and then the dynamic division problem was also studied. Keyvan-Ekbatani et al. [22] studied the feedback gate control method using the simulation network with perimeter gate control and obtained satisfying results with lower total travel time. Aboudolas and Geroliminis [23] used multireservoir networks with well-defined MFDs to design the perimeter and boundary flow control schemes that aimed at distributing the accumulation of vehicles in each reservoir as homogeneously as possible. However, deficiencies are also existed in these control schemes: (i) the change of signal at the cordon may have an influence on the shape of MFD, but the influence was largely ignored; (ii) the common perimeter control needs sufficient road space for stopping and queuing [24, 25], which may not be always allowed. Cordon-based or area-based dynamic congestion pricing at network level has been limited to some simple network due to the complexity of link-based dynamic assignment. Geroliminis and Levinson [26] combined the Vickrey's bottleneck theory with MFD to propose a cordon-based congestion scheme, which is easier implemented in real road networks because of more convenient data collection and much easier computation. Zheng et al. [27] adopted an agent-based simulation to develop and apply a cordon-based dynamic pricing scheme, in which tolls are controlled by MFD. The above works shed light on how to apply MFD to design the dynamic congestion pricing scheme at network level.

This paper aims to combine the MFD with dynamic traffic assignment theory to design a dynamic pricing scheme in heterogeneous networks. A bilevel optimal equilibrium model combining the MFD theory is proposed, which is consistent with the traffic dynamics at network level and can also timely carry out online data analysis and output the expected toll, due to fewer data collection and lower computation cost than link-based equilibrium model. The upper level problem is to minimize the total travel time. The lower level problem is a network equilibrium model with the MFD. In detail, the upper level model is to obtain the optimal toll for the designed system objective from the manager perspective, while the lower level model is to output the expected flow distribution by solving the dynamic user equilibrium assignment with incorporating MFD. Genetic algorithm and method of successive average were used

together to solve the bilevel model. Finally, a numerical test for the two-layer network consisting of the loop expressway network and the linear arterial streets was employed in the simulation environment, which combines the microscopic simulation software VISSIM, mathematic solver software MATLAB, and the development tool Visual Studio as the medium to implement the proposed bilevel toll model.

The remainder of the paper is organized as follows. Section 2 introduces the important features of MFD and explains the reliability for applying MFD to dynamic congestion pricing scheme on the two-layer network. Section 3 describes a bilevel dynamic toll programming model, the solving algorithm, and scheme for model application. In Section 4, the proposed model is tested with the two-layer network in the simulation-feedback environment. Finally, conclusions and recommendations for future work are provided in Section 5.

2. Features of MFD and Research Question

2.1. Features of MFD. Technically, MFD represents traffic characteristics at the network level by aggregating the link flow and density [14, 15]. The weighed space mean density K and the weighed space mean flow Q for a road network with the homogeneous traffic distribution can be expressed as follows:

$$K = \frac{\sum_{\Psi} k_{\Psi} \cdot l_{\Psi} \cdot n_{\Psi}}{\sum_{\Psi} l_{\Psi} \cdot n_{\Psi}} = \frac{N}{L} \quad (1)$$

$$Q = \frac{\sum_{\Psi} q_{\Psi} \cdot l_{\Psi} \cdot n_{\Psi}}{\sum_{\Psi} l_{\Psi} \cdot n_{\Psi}} = \frac{P}{L} \quad (2)$$

where $\Psi \in \Phi$ denotes an individual link in the network links set; k_{Ψ} is the traffic density of link Ψ ; l_{Ψ} denotes the length of link Ψ ; n_{Ψ} is the number of lanes of link Ψ ; q_{Ψ} is the traffic flow of link Ψ ; N called the accumulation is the existing number of vehicles within the current network; P is the network travel production; and L denotes the total length (lane-kilometers) of the network.

The network space mean speed V can be expressed as follows:

$$V = \frac{Q}{K} \quad (3)$$

A representative fundamental diagram relating the space mean flow to space mean density is resemble with the Figure 1.

MFD can be divided into three typical regimes: “free flow” region A with fewer vehicles within the network, which have relative free speed; “maximum flow” region B, where the space mean flow reaches the highest point with the related optimal mean density value called “critical density” that many researchers take advantage of the value to study network control strategies; “congestion flow” region C, where the space mean density is beyond the critical density, the space mean flow decreases with the mean density increasing, and then the congestion will occur. When network is in state D, no one car can nearly arrive to the destination.

Regarding the road network as a reservoir, the state formula of the network system is described by the mass conservation equation as below:

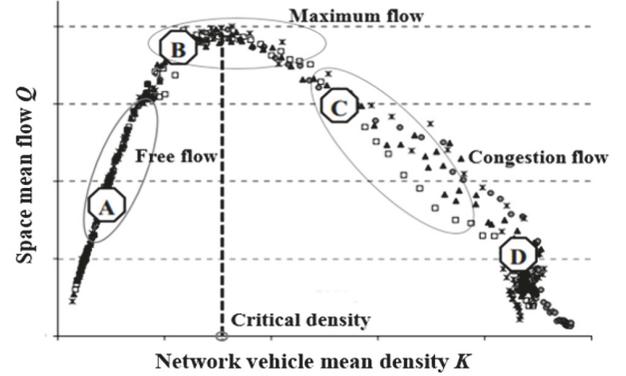


FIGURE 1: Representative MFD (space mean flow versus mean density) (source: reference [14]).

$$\frac{dN(t)}{dt} = q_{in}(t) - q_{out}(K(t)) \quad (4)$$

where $q_{in}(t)$ is the inflow to the network at time t and $q_{out}(K(t))$ is the total outflow representing the trip completion rate at time t . To further analyze the outflow, there is also the unimodal relationship with the space mean density and a robust liner relationship with the space mean flow, which is regarded as another important finding according to field experiment in downtown Yokohama [15].

These formulas provide two important clues: (i) the link-based state formula can be extended to region-based equation thanks to the $q_{out}(t)$ which can be predicted by the state variable $K(t)$; (ii) different from the bottle model that queue will occur when demand goes beyond capacity, the state equation illustrates that traffic system are dynamic, and the average travel time is connected with the initial state of the system as well as input flow.

The above description provides preliminary recognition about the concept of MFD, and features are summarized as follows: (i) MFD is the characteristic of network and not very sensitive to different demand patterns; (ii) the space mean flow can maintain the maximum value when the space mean density value is adjacent to the critical density value; (iii) using existing loop detector data and GPS data can obtain the MFD accurately. Considering these features of MFD, it can be applied to dynamic network cordon pricing.

2.2. Problem Description. Urban transportation network generally can be divided into multiple layers with different properties, such as arterial street, expressway, and metro network [28]. Expressway is a type of road that is built for fast travel with higher design speed [29] and can only be accessed by ramps. Arterial streets are high-capacity urban roads and serve to carry long-distance flows between important centers of activity, which produce large amounts of carbon emission [30]. Compared to the arterial street, expressway has the continuous traffic flow that vehicles do not have to stop for the signal lights in intersections. The expressway is commonly built above the land to decrease land usage and avoid grade crossing in city center region, however, the arterial street is often built on the land for more flexible access. The network

system consisting of the expressway and arterial streets is called the two-layer network.

Drivers often have dilemmas when choosing between the arterial road network and the loop expressway, and the phenomenon exists sometimes that one is congested but the other is in a low density. Dynamic congestion pricing has been proved as an effective strategy to make the users change route by changing the trip utility for the optimal demand distribution on the network. Therefore, there is practical need to develop toll strategy at regional level by means of macroscopic traffic flow theory which is consistent with the network traffic dynamics.

3. Model Formulation

Urban traffic network MFD provides a unimodal low-scatter relationship between space mean density and space mean flow or outflow, if link flows are in roughly homogeneous congestion state within the region. A smoothly decline curve between network mean speed and density can also be identified. The urban network consisting of the multilayers with different properties can be represented by the multi-MFDs respectively. Considering the practical problem, it is essential to take proper toll strategy for balancing the traffic demand in two-layer network consisting of the loop expressway network and arterial streets. The following sections introduce the bilevel model combining MFDs and traffic assignment theory to design a dynamic toll scheme.

3.1. The Lower Level Model. The lower level model combines DUE model with MFD to develop an equilibrium assignment model at aggregated level which regards the network as a special reservoir. Comparing to the traditional link-based dynamic traffic assignment model, the network-based user equilibrium model has some modifications: (i) the trip completion rate is polynomial relationship with the space mean density in network; (ii) the travel time in network depends on the average travel distance and the space mean speed, which can be calculated through the formulation of space mean flow divided by space mean density, and the mean speed is approximately linear relationship with the mean density.

Travel costs include travel time consuming, fuel cost, and road toll. This paper adopts generalized transportation cost function that all travel costs are calculated as equivalent hours using fixed value of time (VOT). Travel costs including travel time and toll are considered as follows:

$$\hat{t}_a(K_a(t), u_a(t)) = t_a(K_a(t)) + \frac{u_a(t)}{VOT} \quad (5)$$

$$t_a(K_a(t)) = \frac{l_a}{V_a(K_a(t))} \quad (6)$$

where $\hat{t}_a(K_a(t), u_a(t))$ is the total travel cost in network a at time t , $u_a(t)$ is the toll in network a at time t , $t_a(K_a(t))$ is the planned travel time in network a at time t ; $K_a(t)$ is space mean density of the network a at time t ; l_a is the average travel distance in network a , $V_a(K_a(t))$ is space mean speed in the network a at time t .

Similar to the formula of link-based dynamic user equilibrium, the objective function is to minimize the integral function representing travel within the two reservoir systems in a time interval T . The MFD-based dynamic user equilibrium model can be formulated as follows:

$$J_l = \min \sum_{a \in A} \int_{i \cdot T}^{(i+1) \cdot T} \int_{K_a(i \cdot T)}^{K_a(t)} \hat{t}_a(w, u_a(t)) dw dt \quad (7)$$

subject to

$$\frac{dN_a(t)}{dt} = \frac{dK_a(t) \cdot L_a}{dt} = q_{in}^a(t) - q_{out}^a(t), \quad \forall a, t \quad (8)$$

$$\sum_{a \in A} q_{in}^a(t) = q_{in}(t), \quad \forall a, t \quad (9)$$

$$q_{out}^a(t) = q_{out}^a(K_a(t)), \quad \forall a, t \quad (10)$$

$$q_{in}^a(t) \geq 0,$$

$$q_{out}^a(t) \geq 0,$$

$$k_a(t) \geq 0,$$

$$\forall a, t$$

where (8) is the network state equation for the network $a \in A$; constraint (9) is the node inflow equilibrium equation; L_a is the total lane length for network a ; $q_{in}(t)$ is the overall traffic demand at time t ; and $q_{in}^a(t)$ is the inflow to network a at time t . Equation (10) is the outflow which is the function of space mean density; constraint (11) ensures that all variables satisfy nonnegative condition.

3.2. The Upper Level Model. The upper level problem is to investigate the desired objective from the manager perspective, in which various purposes may be included within the previous literatures, such as minimizing the total travel time [31, 32]. In the simulation experiment, the total travel time can be formulated as the integral of the number of vehicles (vehicle-hours) in the two-layer network as follows:

$$J_{up} = \min \int_{i \cdot T}^{(i+1) \cdot T} \sum_{a \in A} N_a(t) dt \quad (12)$$

$$u_a(t) \leq u_{max} \quad (13)$$

where $N_a(t)$ is the number of vehicles in network a at time t ; u_{max} is the allowed maximum toll.

3.3. Solution Algorithms. A common characteristic of the model is to determine the optimal value in a set of decision variables, in order to keep the network performance in optimal. Typical examples include network design and road pricing. These questions can be depicted as a Stackelberg game or leader-follower game [33]. The leader refers the traffic manager with the goal of optimizing network usage at system level, and the follower refers the user with optimal

path choice based on route utility. Because the upper and lower problems are mutually restricted and the model is non-convex, it is difficult to solve the bilevel programming model by means of the general optimization method. For the lower model, the model is processed discretely at time dimension. The discrete time form can be considered as a nonlinear programming problem, so the Frank-Wolf algorithm or the method of successive average can be applied to solve the function. From the toll perspective, as time-varying pricing is difficult to be implemented, it is reasonable to adopt constant toll during each evaluation interval.

The above discussion provides the idea that the bilevel model is discrete at time dimension during the evaluation time interval T with the fixed toll. Genetic algorithm with the characteristics of global search is adopted to solve the upper level problem. The lower model is discrete at time dimension. The method of successive average called the approximate equilibrium algorithm is applied to traffic assignment when the calculation accuracy is not strict and time saving is needed. At each iteration, the all-or-nothing assignment is conducted to create auxiliary path flows which are weighted together with the path flows from the former iteration, to obtain path flows for the next iteration. Iterations do not terminate until the algorithm satisfies the convergence criteria and the final solution is approximately equal to the equilibrium solution. These user equilibrium assignment algorithms are common in many research, and the attention point is that, for each time segment, the number of vehicles in each network is so connected with the former stage as the transition process. To simplify the calculation, the instantaneous travel time is assumed to be equal to actual travel time. This study combines Genetic Algorithm with the method of successive average to solve the bilevel programming model, as follows:

Step i. Generate initial population and determining cross-over rate, mutation rate of genetic algorithm, the total number of individuals in each generation, and the maximum evolutionary algebra M . Set evolutionary algebra $m = 0$.

Step ii. For each individual, apply the assignment algorithm to solve the lower level model for each discrete interval, and the state equation $N(j+1) = N(j) + (q_{in}(j) - q_{out}(j)) \cdot \Delta t$ is regarded as the transient process. Then, each responding flow distribution among the group is obtained.

Step iii. Calculate the function value of the upper model corresponding to each toll and obtaining the fitness value. Choose the optimal solution. If $m > M$, output the best solution; otherwise, set $m = m + 1$.

Step iv. Cross over, mutate, and select individuals according to the fitness index value and generate new offspring; then go to Step ii.

3.4. Scheme for Model Application. The proposed model presents the dynamic toll problem as a bilevel problem. The upper level is to minimize total travel time and the lower

level is the dynamic user equilibrium at the network level. Managers provide a charging rate through real-time detection and analysis of road network data, then the users adjust travel routes according to timely utility so that the two alternative networks can maintain the satisfying state. The toll strategy takes full use of the timely traffic information and provides the predicted toll for the given time interval T , with the entire simulation time set as T_a . The timely feedback toll strategy including two stages for the MFD-based bilevel toll model is shown as follows:

First Stage. The two-layer network MFDs were obtained respectively, including the diagrams of network mean flow and mean density, mean speed and mean density, trip complementation rate, and mean density.

Then, the following steps are incorporated as the **second stage** of the model application.

Step i. To simulate for several minutes at start, collect state information for networks, and set $i = 0$.

Step ii. To solve the bilevel programming model to obtain the predictive toll for the next interval according to the timely collected data feedback and MFDs obtained from the first stage, $i = i + 1$.

Step iii. To change the network toll timely and the user will choose the route according to the dynamic utility timely. If $i \cdot T < T_a$, deliver the network traffic data to Step ii; otherwise, stop the simulation experiment and evaluate the result.

4. Case Studies and Results

To test the bilevel model and the corresponding algorithm proposed in the above discussion, a simulation experiment of two-layer network was employed. This study integrated VISSIM, MATLAB, and the development tool Visual Studio as the main control program. In detail, VISSIM and Visual Studio were integrated through COM data communication interface, and Visual Studio and MATLAB were integrated through Dynamic Link Library (DLL). VISSIM was used to simulate traffic flow, MATLAB was used to solve the bilevel programming model, and Visual Studio was used as the intermediate medium.

4.1. MFD for Simulation Network. There are many cities using the circular expressway network to relieve the pressure of transit traffic in the center business region. Paris in France and Chengdu in China are two typical instances as shown in Figures 2(a) and 2(b), which can be summarized as Thomson's urban structure model [34] as shown in Figure 2(c), consisting of grid and ring network. It is suitable to the city center with developed economy and star towns around the perimeter. The simulated two-layer network consists of a loop expressway network and linear arterial streets as shown in Figure 2(d). In the simulation network, the expressway network refers to the bidirectional loop roads including 32 links with two lanes around 500m in length and 3.5m in

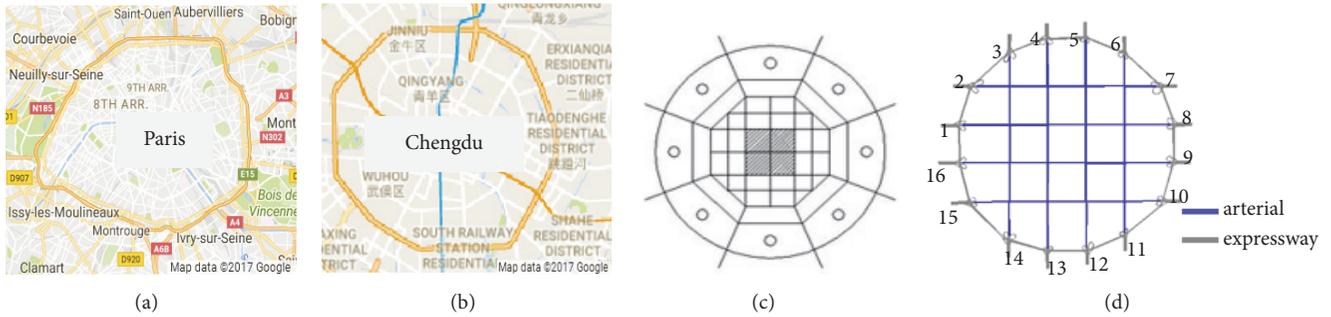


FIGURE 2: Two-layer network layout: (a) city center for Paris, France, (b) city center for Chengdu, China, (c) Thomson urban structure model, and (d) simulation two-layer networks layout.

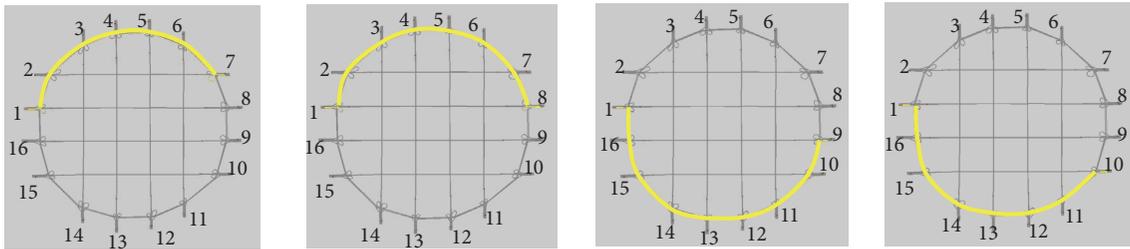


FIGURE 3: Routes in expressway (e.g., Origin 1 with Destinations 7, 8, 9, and 10).

width. The arterial streets network refers to the bidirectional diameter roads including 80 links with two lanes around 500m in length and 3.5m in width. The intersections of arterial streets network adopt pretimed fixed timings, with both Green and Red set as 30s. The free flow speeds within the links of expressway and arterial road are set as 60km/h and 35km/h, respectively.

16 traffic zones are located at the perimeter with the origins and destinations uniformly distributed across the network. To avoid the large variance of trip distance along each network, the OD pairs with the neighbor nodes are not considered. For example, the simulation experiment only considers Origin 1 with Destinations 7, 8, 9, and 10 and Origin 2 with Destinations 8, 9, 10, and 11. As a result, the number of all OD pairs is 64.

Considering the objective is to investigate the influence of the dynamic toll on the route choice between the expressway and arterial streets, so as to achieve a more efficient network performance, e.g., minimizing total travel time, the traffic assigned on routes in each network is simply considered as static instead of dynamic. Many research papers adopted microsimulation that used fixed routing scheme obtained by the static traffic assignment (STA) method [35] to obtain MFD. For example, CORSIM was used in San Francisco, California [14], and VISSIM was used in Orlando, Florida [36], and in Zurich, Switzerland [37]. Recent studies have adopted the dynamic traffic assignment (DTA) modules in VISSIM to obtain the MFD [38]. In dynamic network loading, DTA simulation method is iterative process to search new routes according to the experienced travel time that attempts to provide users equilibrium traffic states. However, DTA method may be not appropriate in this research, which

could induce large variance for the time-varying average trip length in each network due to the fact that some drivers would choose the long detour to avoid the congestion [39]. Moreover, DTA method is more computationally demanding due to a few iterative processes and more unstable as convergence is not always guaranteed. As the research question is to investigate the applicability of dynamic toll at the large network level using the MFD rather than link-based flow-density function, the research simply adopts the STA method with capacity restraint to obtain the fixed routes with the initial traffic demand pattern and then keeps the same OD pattern and multiplies the factors to scale the OD demand for achieving an integrated MFD [37, 40]. The resulting time along each layer is related to the travel distance and network density within the simulation experiment. The average travel distances and the standard deviation (SD) for expressway and arterial streets are shown in Tables 1 and 2, and the typical routes in expressway and arterial streets for Origin 1 with Destinations 7 to 10 are shown in Figures 3 and 4.

The routes of all OD pairs in expressway can be categorized into three types (average trip distance: 3500m, SD: 354m), and the routes of all OD pairs in arterial streets can be categorized into four types (average trip distance: 3030m, SD: 363m). The average trip distance is expected to be constant that the average trip distance is 3.5km for expressway network and 3.0km for arterial streets.

The initial experiment is to obtain the MFDs of the two-layer alternative networks, respectively. The origin-destination (OD) matrix is used to express traffic demand with uniform distribution. The demand factor is used to modify the demand amount for the intact MFD. Detectors are installed at every 100m length in each link and the related data

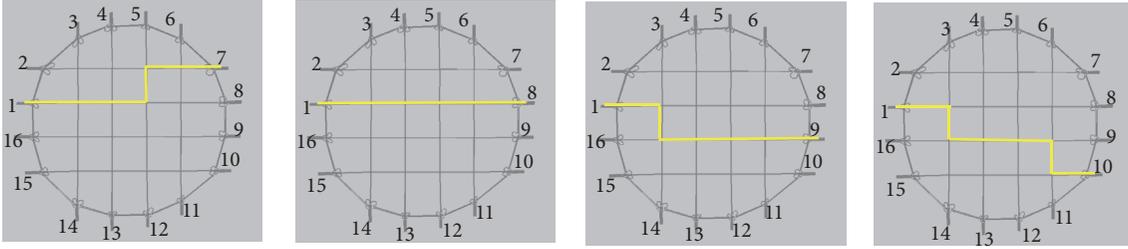


FIGURE 4: Routes in arterial streets (e.g., Origin 1 with Destinations 7, 8, 9, and 10).

TABLE 1: Average trip distance for expressway.

Trip distance (m)	3000	3500	4000
Number of OD pairs	16	32	16
Proportion	25%	50%	25%
Average travel distance (m)	3500 SD = 354		

TABLE 2: Average trip distance for arterial streets.

Trip distance (m)	2500	2750	3250	3500
Number of OD pairs	8	24	16	16
Proportion	12.5%	37.5%	25%	25%
Average travel distance (m)	3030 SD = 363			

including flow and occupancy are collected and aggregated at every 120 seconds. The results are shown in Figure 5.

By adopting least square fitting, six polynomial functions were obtained. For expressway network, space mean flow versus mean density can be expressed as the formula: $Q = 0.005K^3 - 1.337K^2 + 88.923K (R^2 = 0.96)$, space mean speed versus mean density can be expressed as the formula: $V = -0.98K + 84.92 (R^2 = 0.97)$, and network outflow versus mean density can be expressed as the formula: $q_{out} = 0.059K^3 - 13.35K^2 + 751.76K (R^2 = 0.97)$. For arterial network, space mean flow versus mean density can be expressed as the formula: $Q = 0.0062K^3 - 1.15K^2 + 49.76K (R^2 = 0.98)$, space mean speed versus mean density can be expressed as the formula: $V = -0.51K + 36 (R^2 = 0.90)$, network outflow versus mean density can be expressed as the formula: $q_{out} = 0.027K^3 - 11.25K^2 + 634K (R^2 = 0.91)$.

4.2. Dynamic Pricing Experiment Evaluation. As indicated in Figures 5(a), 5(b), 5(d), and 5(e), the critical densities for expressway network and arterial streets are both about 40veh/km, and the network mean speeds for expressway network and arterial streets at the critical density are around 40km/h and 20km/h, respectively. According to user equilibrium principle, due to higher mean speed for expressway network even though in right band of MFD which is larger than the critical density, when in high traffic demand condition, the expressway network would operate in congestion flow state which is easy to cause congestion; however, the arterial streets would operate in the free flow state. The purpose of the simulation experiment is to decrease the density of expressway network to maintain the two-layer density load

TABLE 3: Improvement with applying the toll scheme.

Index	No toll	Toll	Improvement
Total Travel Time (h)	6415.6	5080.8	1334.8
Total Collected Revenue (\$)	-	13620.2	13620.2
Average Speed (km/h)	33.2	42.3	9.1
Average Travel Cost (h)	0.123	0.114	0.009

in an optimal state. Consequently, only the freeway network is charged with extra toll in the simulation experiment considering the significant difference of average speeds between the two-layer networks, however, if in field experiment, the two-layer networks would cooperatively take toll schemes. In this research, deterministic and homogeneous population characteristics are assumed with fixed VOT 16\$/h [41, 42], and the maximum toll is set as 10\$ [43].

The fluctuated demand profile including two peak periods is presented in Figure 6.

In this section, simulation experiments for applying both the dynamic toll strategy and no toll were carried out. The simulation for first 10 minutes is to produce preliminary volume in the network, and the data about network state are collected and aggregated for each 2 minutes. Table 3 presents the improvements after applying the toll scheme.

The total travel time savings and the total collected revenue were investigated. The total travel time decreases from 6415.6h to 5080.8h. The total collected revenue is 13620.2\$. Considering the value of time is 16\$/h, the total toll paid in the study network is smaller than the total time savings, by 36.2%. The time savings are much higher than the toll paid, which is a promising result at an aggregated level. The average travel cost without toll is 0.123h, and the average travel cost with toll represented by time cost decreased to 0.114h. Moreover, the average speed increased by 9.1km/h.

The time-dependent tolls are presented in Figure 7, while the average speeds with toll or not were compared in Figure 8. It is found that the time-dependent toll diagram also has two peaks, and the highest toll values are 0.66\$ per trip and 0.69\$ per trip in the fourth and ninth periods, respectively. For the two high peak demand periods, the average speeds with no toll are only 29.5 km/h and 26.1 km/h, while the average speeds with toll are 39 km/h and 38.3 km/h. These further illustrate that the toll scheme is effective in improving the speed for the two-layer network. Moreover, although tolls at all time are above some base level of about 0.45\$, under

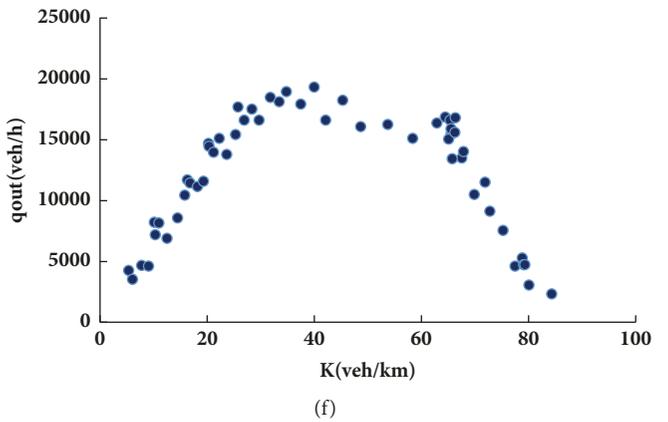
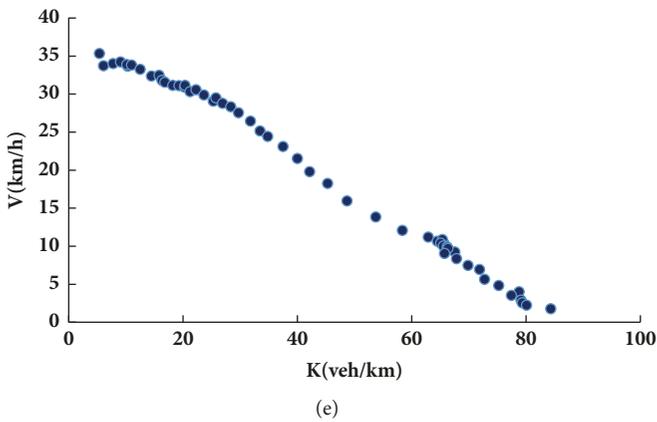
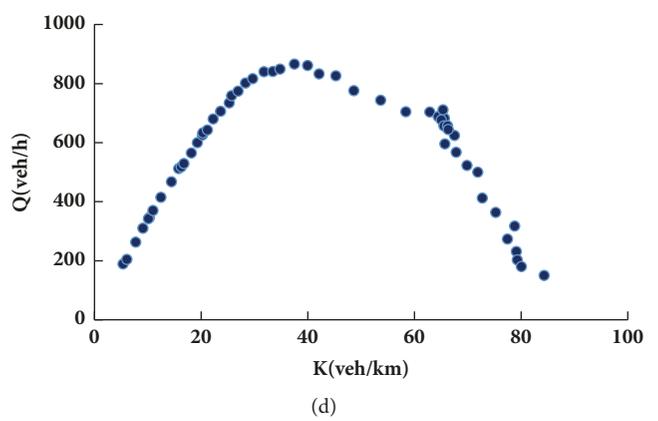
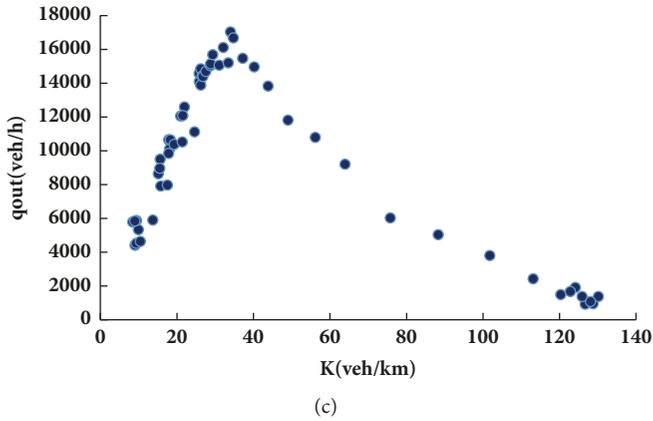
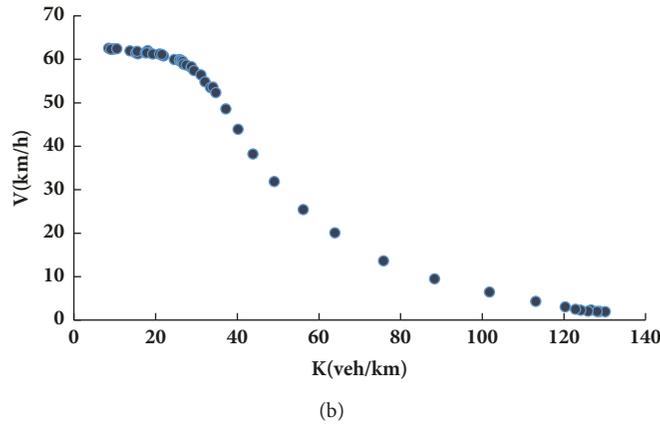
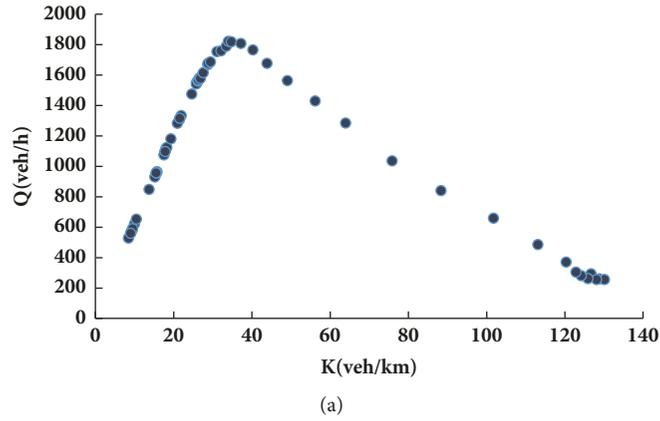


FIGURE 5: MFD of the two-layer network: (a) expressway network: space mean flow versus mean density, (b) expressway network: space mean speed versus mean density, (c) expressway network: network outflow versus mean density, (d) arterial streets: space mean flow versus mean density, (e) arterial streets: space mean speed versus mean density, and (f) arterial streets: network outflow versus mean density.

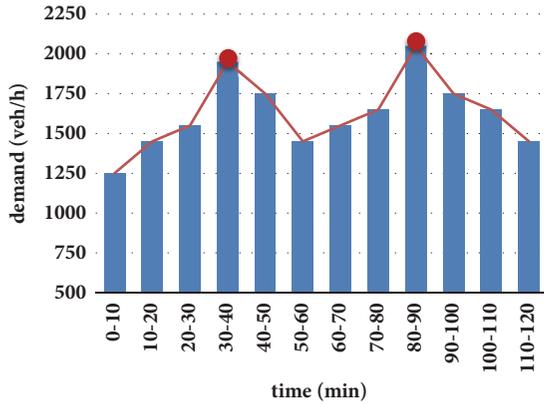


FIGURE 6: Fluctuated demand profile including two peak periods.

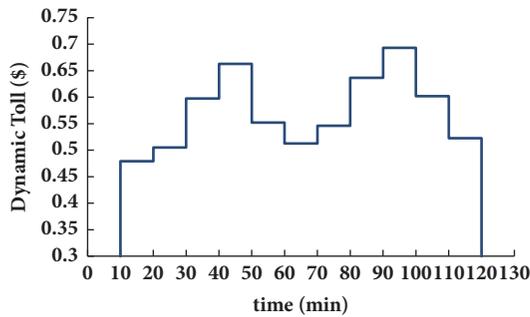


FIGURE 7: Time-dependent toll.

bottleneck models and departure time choice, the criterion value may tend to be null as the base level.

Furthermore, Figures 9(a) and 9(b) present the detailed MFD for expressway and arterial streets in toll and no toll conditions. Under no toll condition, more drivers would choose expressway for the higher travel speed, then the average density is far beyond the critical density (Figure 9(a)), which is much easier to cause traffic congestion. However, the average density of expressway falls on the left part in toll. For arterial streets, the range of the average density is from 5 to 35veh/km, while it mainly falls between 15 to 25veh/km under toll condition (Figure 9(b)), which remains a larger reservoir capacity for more traffic demand. It can be explained that during toll condition the road networks operate in much more optimal state by analyzing the MFD.

5. Conclusions

In this paper, a dynamic toll strategy at a two-layer network is proposed. The idea is to adopt a macroscopic traffic model, i.e., the macroscopic fundamental diagram (MFD), to determine the optimal toll. The bilevel programming model combining the traffic assignment principle and the theory of MFD is proposed. The upper level model is to decrease the total travel time in the two-layer network by optimizing the toll and the lower level model extends the link-based dynamic user equilibrium to network-based dynamic user equilibrium by means of the physics of network traffic expressed by

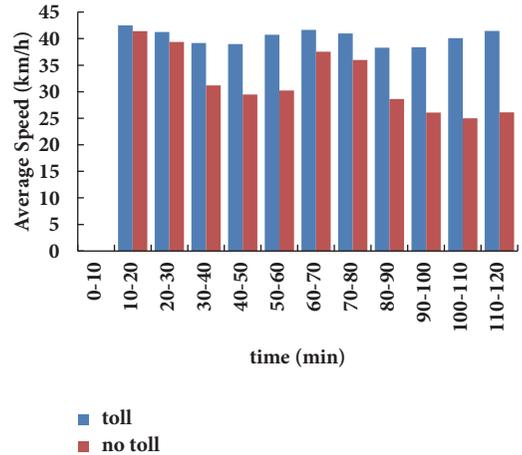


FIGURE 8: Average speed in toll and no toll condition.

the MFD. Genetic algorithm and the method of successive average are applied for the solution to the discretized bilevel model. The integration of VISSIM, MATLAB, and Visual Studio as the main control program is proposed to validate the proposed bilevel toll model for the given two-layer network.

The results indicate that the real-time toll strategy is effective for keeping the network operating in a better state. On other words, the arterial road resources can be utilized in a more effective way and the expressway network becomes more moderately to avoid the road congestions. Compared with no toll condition, in which UE criterion depends just on the average travel time, the bilevel toll model can significantly decrease the total travel time. Because the MFD-based dynamic toll model at aggregated level can overcome the calculation complexity of link-based assignment and the data collection and statistics to evaluate the network state are easier, the proposed dynamic toll strategy is anticipated to become more effective in a larger urban network.

The proposed dynamic tolling is promising to solve the exist problem in charging toll at network aggregated level for the optimal traffic distribution on the overall urban network. However, some problems need more efforts to solve in future. Users with different value of time (VOT) and heterogeneous travel distances should be thought, and the agent-based simulation can also be used to implement the toll model [44–46]. Moreover, the multiple modes traffic should be incorporated considering the toll is also to encourage users to take the public transport. To design more practical management scheme for the multi-MFDs regions, it is especially necessary to develop effective algorithms to decompose the large-scale center network into interconnected subdomains represented by the own MFD, such as the shortest distance decomposition algorithm for partitioning large-scale network [47].

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

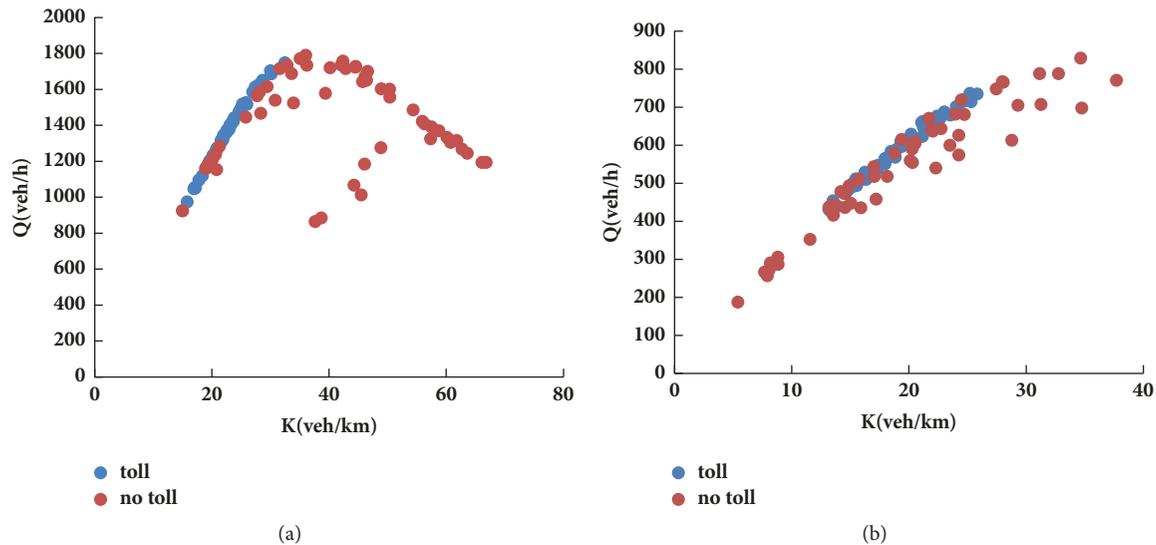


FIGURE 9: MFD in toll and no toll conditions: (a) for expressway; (b) for arterial streets.

Disclosure

Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to express their appreciation to Dr. Zhong-Ren Peng from Center for UAV Applications and ITS Research, School of Naval Architecture Ocean and Civil Engineering, Shanghai Jiao Tong University, for his valuable suggestions and assistance in this study. The research was supported in part by the Major Project of National Social Science Foundation of China [16ZDA048], the Humanities and Social Science Research Project, Ministry of Education, China [18YJCZH011], and the Shanghai Municipal Natural Science Foundation [17ZR1445500], China.

References

- [1] S.-Y. Phang and R. S. Toh, "From manual to electronic road congestion pricing: The Singapore experience and experiment," *Transportation Research Part E: Logistics and Transportation Review*, vol. 33, no. 2, pp. 97–106, 1997.
- [2] A. A. Walters, "The theory and measurement of private and social cost of highway congestion," *Econometrica*, vol. 29, no. 4, pp. 676–699, 1961.
- [3] W. Vickrey, "Pricing and resource allocation in transportation and public utilities," *American Economic Review*, vol. 53, no. 2, pp. 452–465, 1963.
- [4] Y. Sheffi, "Urban transportation networks: equilibrium analysis with mathematical programming methods," *Transportation Research Part A: General*, vol. 20, no. 1, pp. 76–77, 1986.
- [5] P. Ferrari, "Road pricing and network equilibrium," *Transportation Research Part B: Methodological*, vol. 29, no. 5, pp. 357–372, 1995.
- [6] F. Leurent, "Cost versus time equilibrium over a network," *European Journal of Operational Research*, vol. 71, no. 2, pp. 205–221, 1993.
- [7] H. Yan and W. H. K. Lam, "Optimal road tolls under conditions of queueing and congestion," *Transportation Research Part A: Policy and Practice*, vol. 30, no. 5, pp. 319–332, 1996.
- [8] H. Yang and M. G. H. Bell, "Traffic restraint, road pricing and network equilibrium," *Transportation Research Part B: Methodological*, vol. 31, no. 4, pp. 303–314, 1997.
- [9] W. Vickrey, "Congestion theory and transport investment," *American Economic Review*, vol. 59, pp. 251–261, 1969.
- [10] H. Yang and Q. Meng, "Departure time, route choice and congestion toll in a queueing network with elastic demand," *Transportation Research Part B: Methodological*, vol. 32, no. 4, pp. 247–260, 1998.
- [11] M. Carey and D. Watling, "Dynamic traffic assignment approximating the kinematic wave model: System optimum, marginal costs, externalities and tolls," *Transportation Research Part B: Methodological*, vol. 46, no. 5, pp. 634–648, 2012.
- [12] X. Ban and H. X. Liu, "A link-node discrete-time dynamic second best toll pricing model with a relaxation solution algorithm," *Networks and Spatial Economics*, vol. 9, no. 2, pp. 243–267, 2009.
- [13] J. A. Laval and F. Castrillón, "Stochastic approximations for the macroscopic fundamental diagram of urban networks," *Transportation Research Part B: Methodological*, vol. 81, pp. 904–916, 2015.
- [14] N. Geroliminis and C. F. Daganzo, "Macroscopic modeling of traffic in cities," in *Transportation Research Board 86th annual meeting (No. 07-0413)*, Washington D.C., Jan 13–17, 2007.
- [15] N. Geroliminis and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: some experimental findings," *Transportation Research Part B: Methodological*, vol. 42, no. 9, pp. 759–770, 2008.

- [16] L. Leclercq and N. Geroliminis, "Estimating MFDs in simple networks with route choice," *Procedia-Social and Behavioral Sciences*, vol. 80, pp. 99–118, 2013.
- [17] N. Geroliminis and J. Sun, "Properties of a well-defined macroscopic fundamental diagram for urban traffic," *Transportation Research Part B: Methodological*, vol. 45, no. 3, pp. 605–617, 2011.
- [18] N. Geroliminis, J. Danés, and M. Á. Estrada, "Multimodal Macroscopic Fundamental Diagram for "Car-Bus" Mixed Traffic Signalized Corridors: Application in City of Barcelona," in *The Transportation Research Board 92nd Annual Meeting (No. 13-1089)*, Mixed Traffic Signalized Corridors, Washington D.C., Jan 13-17, 2013.
- [19] Y. Ji and N. Geroliminis, "On the spatial partitioning of urban transportation networks," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1639–1656, 2012.
- [20] D. Sun, K. Zhang, and S. Shen, "Analyzing spatiotemporal traffic line source emissions based on massive didi online car-hailing service data," *Transportation Research Part D: Transport and Environment*, vol. 62, pp. 699–714, 2018.
- [21] K. Zhang, D. J. Sun, S. Shen, and Y. Zhu, "Analyzing spatiotemporal congestion pattern on urban roads based on taxi GPS data," *Journal of Transport and Land Use*, vol. 10, no. 1, pp. 675–694, 2017.
- [22] M. Keyvan-Ekbatani, A. Kouvelas, I. Papamichail, and M. Papageorgiou, "Exploiting the fundamental diagram of urban networks for feedback-based gating," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1393–1403, 2012.
- [23] K. Aboudolas and N. Geroliminis, "Perimeter and boundary flow control in multi-reservoir heterogeneous networks," *Transportation Research Part B: Methodological*, vol. 55, pp. 265–281, 2013.
- [24] S. Chen and D. J. Sun, "An Improved Adaptive Signal Control Method for Isolated Signalized Intersection Based on Dynamic Programming," *IEEE Intelligent Transportation Systems Magazine*, vol. 8, no. 4, pp. 4–14, 2016.
- [25] Y. Zhang, D. J. Sun, and A. Kondyli, "An empirical framework for intersection optimization based on uniform design," *Journal of Advanced Transportation*, vol. 2017, 2017.
- [26] N. Geroliminis and D. M. Levinson, "Cordon pricing consistent with the physics of overcrowding," in *Transportation and Traffic Theory*, pp. 219–240, Springer, Golden Jubilee, 2009.
- [27] N. Zheng, R. A. Waraich, K. W. Axhausen, and N. Geroliminis, "A dynamic cordon pricing scheme combining the Macroscopic Fundamental Diagram and an agent-based traffic model," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 8, pp. 1291–1303, 2012.
- [28] D. J. Sun and S. Guan, "Measuring vulnerability of urban metro network from line operation perspective," *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 348–359, 2016.
- [29] H. Liu, L. Zhang, D. Sun, and D. Wang, "Optimize the Settings of Variable Speed Limit System to Improve the Performance of Freeway Traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 3249–3257, 2015.
- [30] D. Sun, Y. Zhang, R. Xue, and Y. Zhang, "Modeling carbon emissions from urban traffic system using mobile monitoring," *Science of the Total Environment*, vol. 599–600, pp. 944–951, 2017.
- [31] H. Yang and H. J. Huang, *Mathematical and Economic Theory of Road Pricing*, Elsevier, Oxford, UK, 2005.
- [32] J. A. Laval, H. W. Cho, J. C. Muñoz, and Y. Yin, "Real-time congestion pricing strategies for toll facilities," *Transportation Research Part B: Methodological*, vol. 71, pp. 19–31, 2015.
- [33] A. H. Meimand, *Differential Stackelberg Games and their Application to Dynamic Pricing, Production Planning Network Design, and Logistics [Ph.D. thesis]*, The Pennsylvania State University, 2013.
- [34] J. M. Thomson, "Toward better urban transport planning in developing countries," *World Bank Staff Working Paper*, vol. 600, 1983.
- [35] X.-Y. Ni, D. Sun, and Z.-R. Peng, "An improved incremental assignment model for parking variable message sign location problem," *Journal of Advanced Transportation*, vol. 49, no. 7, pp. 817–828, 2015.
- [36] V. Gayah and V. Dixit, "Using mobile probe data and the macroscopic fundamental diagram to estimate network densities," *Transportation Research Record*, no. 2390, pp. 76–86, 2013.
- [37] J. Ortigosa, M. Menendez, and H. Tapia, "Study on the number and location of measurement points for an MFD perimeter control scheme: a case study of Zurich," *Journal on Transportation Logistics*, vol. 3, no. 3, pp. 245–266, 2014.
- [38] J. Ortigosa, M. Menendez, and V. V. Gayah, "Analysis of network exit functions for various urban grid network configurations," *Transportation Research Record*, vol. 2491, pp. 12–21, 2015.
- [39] L. Leclercq, C. Parzani, V. L. Knoop, J. Amourette, and S. P. Hoogendoorn, "Macroscopic traffic dynamics with heterogeneous route patterns," *Transportation Research Part C: Emerging Technologies*, vol. 59, pp. 292–307, 2015.
- [40] C. F. Daganzo, V. V. Gayah, and E. J. Gonzales, "Macroscopic relations of urban traffic variables: bifurcations, multivaluedness and instability," *Transportation Research Part B: Methodological*, vol. 45, no. 1, pp. 278–288, 2011.
- [41] N. Zheng, G. Rerat, and N. Geroliminis, "A time-dependent area-based pricing scheme for multimodal urban networks with user adaptation: An agent-based approach," in *Proceedings of the 2014 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014*, pp. 471–476, China, October 2014.
- [42] K. W. Axhausen, S. Hess, A. Koenig, J. Bates, M. Bierlaire, and G. Abay, "State-of-the-Art Estimates of Swiss Value of Travel Time Savings. the Transportation Research Board 86th Annual Meeting," *Washington D.C.*, 2007.
- [43] Z. Liu, Q. Meng, and S. Wang, "Speed-based toll design for cordon-based congestion pricing scheme," *Transportation Research Part C: Emerging Technologies*, vol. 31, pp. 83–98, 2013.
- [44] Y. Huang, D. Sun, and L. Zhang, "Identifying Intrapersonal Heterogeneity of Driver Behavior from Taxi GPS Traces," *Accident Analysis & Prevention*, vol. 117, pp. 318–327, 2018.
- [45] X. Ni and D. Sun, "Agent-based modelling and simulation to assess the impact of parking reservation system," in *Journal of Advanced Transportation*, vol. 37, pp. 1–10, 2017.
- [46] D. Sun, X. Ni, and L. Zhang, "A discriminated release strategy for parking variable message sign display problem using agent-based simulation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 1, pp. 38–47, 2015.
- [47] P. Johnson, D. Nguyen, and M. Ng, "Large-scale network partitioning for decentralized traffic management and other transportation applications," *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, vol. 20, no. 5, pp. 461–473, 2016.

