Research Article

Global Optimization for Transport Network Expansion and Signal Setting

Haoxiang Liu,¹ David Z. W. Wang,¹ and Hao Yue²

¹*School of Civil and Environmental Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798*

²*MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing 100044, China*

Correspondence should be addressed to David Z. W. Wang; wangzhiwei@ntu.edu.sg

Received 15 June 2015; Revised 22 September 2015; Accepted 29 September 2015

Academic Editor: Luca D'Acierno

Copyright © 2015 Haoxiang Liu et al. 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.

This paper proposes a model to address an urban transport planning problem involving combined network design and signal setting in a saturated network. Conventional transport planning models usually deal with the network design problem and signal setting problem separately. However, the fact that network capacity design and capacity allocation determined by network signal setting combine to govern the transport network performance requires the optimal transport planning to consider the two problems simultaneously. In this study, a combined network capacity expansion and signal setting model with consideration of vehicle queuing on approaching legs of intersection is developed to consider their mutual interactions so that best transport network performance can be guaranteed. We formulate the model as a bilevel program and design an approximated global optimization solution method based on mixed-integer linearization approach to solve the problem, which is inherently nonlinear and nonconvex. Numerical experiments are conducted to demonstrate the model application and the efficiency of solution algorithm.

1. Introduction

The growth of urban population, social progress, and economic prosperity lead to an increasing travel demand. Therefore, it is imperative for the planners to understand how to determine the optimal transportation planning to ensure best transportation system performance so that the fast growing travel demand can be addressed in the planned transport network. One important issue in transportation planning is to find out the optimal decision on road capacity design, that is, the optimal road capacity expansions and optimal planning on new links addition. Another intrinsic issue to be addressed in urban transportation planning is the equilibrium network signal setting, which seeks to optimize the road capacity allocation in the network by devising the signal control strategy.

In the literature, network design problem (NDP) has received extensive research attention for decades [1–5]. Historically, the NDP has emerged to cater for the faster growing

travel demand via the optimal decision on the expansion of urban road network, subject to the limited resources available for system improvement. The objective of the NDP is to optimize a specific performance index of the transportation system (e.g., minimize the total monetary travel cost), accounting for drivers' routing behavior at the same time. There are mainly three types of the NDP: one that is dealing with the continuous capacity expansion of links, referred to as continuous NDP (CNDP) [1, 2, 6], another one that determines the optimal decision of the new links addition, referred to as discrete NDP (DNDP) [7–9], and a combination of the former two types, referred to as mixed NDP (MNDP) [10]. Recently, some researchers also considered travel demands variation over planning horizon in NDP, which is categorized into a time-dependent NDP [11, 12]. The NDP is often formulated as a bilevel problem with upper-level describing the expansion decision-making of transport planners to minimize the investment and travel cost, while lower-level is representing the traffic assignment on the network for given expansion

plan. For comprehensive review on models and algorithms of the NDP, one can refer to Yang and Bell [13] and Farahani et al. [14]. The equilibrium network signal setting (ENSS) problem involves the optimal signal control of the transport system to improve the network performance (e.g., minimize delay, fuel consumption, or travel time). Studies on ENSS include two categories: local optimization of signal setting for individual intersection and global signal setting for the network. Local methods are widely studied in previous studies [15–18], which assume fixed and exogenously given network flows. On the other hand, global methods [19–22] determine signal setting parameters and endogenous equilibrium network flows simultaneously.

The solution of both the CNDP and ENSS problems is a challenging issue because the model formulations are intrinsically nonlinear and nonconvex, which makes it difficult to even find solution algorithm for local optimum [23, 24]. Sheffi and Powell [20] proposed a gradient projection method on solving the ENSS problem and it is subsequently extended in the work of Cascetta et al. [25], Lee and Machemehl [26], and Cipriani and Fusco [27]. Chiou [28] exploited another descent approach implementing four variants of gradient-based methods to solve the general CNDP. In addition, based on the results of Tobin and Friesz [29], a number of sensitivity analysis based algorithms were introduced to solve bilevel transport planning and management problems: Friesz et al. [4], Sumalee et al. [30], and Connors et al. [31] applied this method to the general CNDP and relevant problems; Yang and Yagar [22] developed an algorithm to solve the ENSS with capacity constraints; Ying et al. [32] further developed a sensitivity analysis based algorithm on ENSS, which significantly improved computing efficiency by avoiding the path-enumerating process. Moreover, Meng et al. [33] imposed an augmented Lagrangian method to solve an equivalent single-level problem of CNDP with nonconvex and continuously differentiable properties, which was transformed via a marginal function tool from the bilevel program. However, for the nonconvex nonlinear CNDP and the ENSS problems, most of the gradient-based algorithms can only obtain local optimal solutions and the solutions are heavily dependent on the selection of initial points. Besides, getting gradient information will be much more difficult with the increasing size of the network. The other type of solution method is metaheuristic algorithm. Abdulaal and LeBlanc [1] suggested Hooke-Jeeves algorithm to solve the CNDP by transferring the CNDP into an unconstrained nonlinear program. Then a simple metaheuristic algorithm called equilibrium decomposed optimization method was developed by Suwansirikul et al. [6] for practical use. It increased the computing efficiency through decomposing the CNDP into interacting subproblems and calculating user equilibrium at each iteration. Besides, genetic algorithm [24, 26, 34, 35], simulated annealing [36, 37], and tabu search [18, 38, 39] are also applied to the CNDP and ENSS problem. The advantage of the metaheuristic algorithms is that they are simple and can be easily implemented on the nonconvex nonlinear problems with many local optimal solutions. What is more, as derivative information is not needed, they avoid the troublesome gradient calculation process [35]. However, generally long

computational time is needed in the stochastic global search, and the global optimal solution cannot be guaranteed. Recently, Wang and Lo [40] proposed a method to obtain the global optimal solution of linearized CNDP. This method is soon generalized in the work of Luatthep et al. [10] for solving the mixed NDP. Many other research works developed global optimization solution algorithms to solve the network design problems [41–43]. Szeto et al. [44] applied artificial bee colony algorithm to solve network design problem.

Basically, most of the previous research works investigated the two planning problems, road capacity design in NDP and equilibrium network signal setting (ENSS), separately. However, in practice, NDP determines the link capacity, while ENSS determines the capacity allocation; and the combination of the road capacity design and the capacity allocation design is needed that fully determines the resultant equilibrium travel pattern and the transportation network performance. Most of previous studies addressed the two problems separately, regarding the other problem as exogenously given, which ignores the mutual interaction between them in determining the transport network performance. Ziyu and Yifan [45] developed an integrated model to combine the concept of reserve capacity and CNDP, which aimed to maximize the reserve capacity of the network. A heuristic algorithm based on sensitivity analysis was applied to solve their problem. Cantarella et al. [18] used LOSS (local optimization of signal setting) method to study the discrete network design problem (lane layout) combined with signal setting and proposed several heuristic methods to solve it. Nevertheless, the presented solution algorithms cannot guarantee the global optimization solutions of the two problems, both of which did not notice the queuing phenomenon on an approaching leg of intersection. In this paper, we would present a combined network design and signal setting problem (CNDSSP) with consideration of queuing vehicles storage constraint on approaching legs of intersections to minimize the total transportation system cost. The problem is formulated as a nonlinear and nonconvex model. Moreover, we design an approximated global optimization solution method to solve the formulated model, which if solved by a local optimal algorithm may result in urban transportation plan that is deviated from the true optimal solution. Following the approximated global optimization solution idea introduced in the work of Wang and Lo [40], we transform the originally formulated bilevel program into a single-level mixed-integer linear program by applying a mixed-integer linearization approach. Noting that the solution efficiency of this method is highly restricted by the number of introduced integer variables, we apply a new mixed-integer linearization model, which significantly reduces the number of integer variables needed and thus improves the solution efficiency. Specifically, the number of integer variables introduced for the linearization follows logarithmic relationship with the size of the partition scheme, which requires much smaller number of integer variables even if high linearization accuracy and refined partition scheme are needed. In summary, this study contributes to literature in two aspects: firstly, we propose a combined network design and signal setting transport planning model with consideration of queuing vehicle storage

restriction on approaching legs of intersections to take into account their interaction and ensure the true optimal transportation planning. Secondly, an efficient global optimization solution method is developed to seek the approximated globally optimal design of the transportation planning problem. Numerical studies are conducted to demonstrate the solution quality by comparing the proposed model and solution algorithm with the other existing models and solution methods.

The remainder of this paper is organized as follows. In Section 2, the bilevel model of combined network design and signal setting is stated and then reformulated as a single-level mixed-integer program (MIP). Section 3 demonstrates the single-level model linearization process. A numerical example is given in Section 4 to illustrate the proposed model and solution method and to compare the computational results with other models and algorithms. Some final remarks are presented in Section 5.

2. Model Formulation

The following notation is used in the formulation.

Parameters

- A^0 : the set of links in the network,
- A : the set of links with signal control, $A \subset A^0$,
- $\bar{A}$: the set of links without signal control, $\bar{A} \subset A^0$,
- A_i : the set of approaching links for intersection $i \in I$ in the network,
- I : the set of signal-controlled intersections in the network,
- W : the set of origin-destination (OD) pairs,
- R_w : the set of paths between OD pair $w \in W$,
- q_w : the fixed OD travel demand, $\forall w \in W$,
- Δ : the link-path incidence matrix, $\Delta = [\delta_{ap}^w]$, $\forall w \in W$, $a \in A^0$, $p \in R_w$, where $\delta_{ap}^w = 1$ if link a belongs to path p between OD pair w , and $\delta_{ap}^w = 0$ otherwise,
- ρ : the conversion factor which turns travel time into monetary cost.

Variables

- x_a : the traffic flow of link $a \in A^0$,
- λ_a : the green split of link $a \in A$; $\lambda_a = 1$ if $a \in \bar{A}$,
- d_a : the queuing delay on link $a \in A$; $d_a = 0$ if $a \in \bar{A}$,
- S_a : the capacity of link $a \in A^0$ after expansion,
- f_p^w : path flow of path p between OD pair $w \in W$,
- c_p^w : travel time of path p between OD pair $w \in W$,
- π^w : minimum path travel cost between OD pair $w \in W$.

Functions

- $t_a(x_a, \lambda_a, S_a)$: travel time function of link $a \in A^0$, which is the function of link flow x_a , green split λ_a , and link capacity S_a ,

$g_a(S_a)$: the link expansion cost function, $\forall a \in A^0$, which is assumed to be the function of S_a ,

$Q_a(S_a)$: the link storage capacity function, $\forall a \in A^0$, which is assumed to be the function of S_a .

2.1. Bilevel Model for CNDSSP. Given the OD demand pattern, which is supposed to be relatively stable within the planning horizon and current network layout, the combined continuous network design and signal control problem in this study aims to determine which links should be expanded, how much the link capacity should be improved, and what the green split of each signal-controlled link is. The objective is to minimize the total travel time cost and the investment on link capacity enhancement subject to the feasible range of signal settings and constraints of maximum link capacity expansions.

The combined network design and signal setting problem can be modeled into a bilevel program stated as follows:

$$\begin{aligned} \text{Minimize} \quad & F(\lambda, S, x(\lambda, S), d(\lambda, S)) \\ & = \sum_{a \in A^0} \rho x_a \cdot (t_a(x_a, \lambda_a, S_a) + d_a) \\ & \quad + \sum_{a \in A^0} g_a(S_a) \end{aligned} \quad (1)$$

$$\begin{aligned} \text{Subject to} \quad & \lambda_a^{\min} \leq \lambda_a \leq \lambda_a^{\max}, \quad a \in A, \\ & \sum_{a \in A_i} \lambda_a = 1, \quad i \in I, \end{aligned} \quad (2)$$

$$S_a^0 \leq S_a \leq S_a^{\max}, \quad a \in A^0, \quad (3)$$

$$d_a(\lambda, S) \cdot x_a(\lambda, S) \leq Q_a(S_a), \quad a \in A, \quad (4)$$

while the equilibrium traffic flow x is determined as follows:

$$\text{Minimize} \quad Z(x, f) = \sum_{a \in A^0} \int_0^{x_a} t_a(v, \lambda_a, S_a) dv \quad (5)$$

$$\begin{aligned} \text{Subject to} \quad & \sum_{w \in W} \sum_{p \in R_w} f_p^w \delta_{ap}^w = x_a, \quad a \in A^0, \\ & \sum_{p \in R_w} f_p^w = q^w, \quad \forall w \in W, \\ & f_p^w \geq 0, \quad \forall p \in R_w, \quad \forall w \in W, \\ & x_a \leq \lambda_a S_a, \quad a \in A, \end{aligned} \quad (6)$$

where $\lambda_a^{\max}$ and $\lambda_a^{\min}$ are the upper and lower bound of green split of link $a \in A$, S_a^0 is the current capacity of link $a \in A$, and $S_a^{\max}$ is the maximum link capacity after expansion due to the restriction of urban land use planning.

In the above formulation, the upper-level of the model is to seek the optimal green splits and road capacity enhancement to minimize the total travel time cost and investment on link capacity expansion. Here, total travel time cost includes all flow-dependent running time, signal delay, and queuing

delay at link exits due to road saturation in the network. As was done in the work of Yang and Yagar [22], the running time and signal delay are captured by travel time function $t_a(x_a, \lambda_a)$ and the queuing delay is endogenously solved and denoted as d_a . Constraint (2) states the given green split restriction and entails the conservation equation depending on the specific signal phase structure of an intersection. Equation (3) describes the boundary constraints of link capacity expansion. As a reasonable signal control should consider the queuing length on the approaching leg of an intersection to avoid blocking the upstream section, (4) enforces that number of queuing vehicles on an approaching leg should be restricted to the link storage capacity. x_a and d_a can be endogenously obtained by solving the lower-level program, which determines the equilibrium travel pattern following the deterministic user equilibrium (UE) principle. Constraint (6) ensures that the total travel demand and link flow conservation are fulfilled. Equation (7) denotes link exit capacity constraints. The Lagrange multipliers associated with the capacity constraints are exactly the corresponding queuing delay d_a due to limited link capacity. For any given signal settings and link capacity, drivers' route choice behavior and the resultant equilibrium travel pattern can be predicted through the lower-level program with the assumption that all drivers perfectly know all the travel costs and queuing delay information of the whole network. Note that the queuing delay d_a is an implicit variable in the bilevel model.

2.2. Single-Level MIP for CNDSSP. In this section, we transform the bilevel problem into a single-level program, which is then linearized into a mixed-integer linear program (MILP), whose global optimality can be guaranteed.

The UE condition assumes that each driver has perfect information of travel time on each possible path and chooses the route with minimum travel time; that is, the travel times on all the routes used are equal and less than those experienced on any unused routes [46]. It implies that the following relationship should be satisfied:

$$\begin{aligned} f_p^w &\geq 0 \implies \\ c_p^w - \pi^w &= 0, \\ f_p^w &= 0 \implies \\ c_p^w - \pi^w &\geq 0, \end{aligned} \quad (8)$$

$$\forall p \in R_w, \quad w \in W,$$

wherein $c_p^w = \sum_{a \in A^0} \delta_{ap}^w (t_a(x_a, \lambda_a, S_a) + d_a)$, which is the total travel time cost for individual traveler, including both d_a the queuing delay and t_a the running time and signal delay.

The above UE condition can be expressed into the form of complementary condition:

$$\begin{aligned} f_p^w \cdot (c_p^w - \pi^w) &= 0, \\ c_p^w - \pi^w &\geq 0, \end{aligned} \quad (9)$$

$$\forall p \in R_w, \quad w \in W.$$

Next, a set of mixed-integer constraints can be introduced to substitute (9). Referring to Lo [47, 48] and Wang and Lo [40], the equivalent set of mixed-integer constraints are stated as follows:

$$\begin{aligned} L \cdot \sigma_p^w + \varepsilon &\leq f_p^w \leq U \cdot (1 - \sigma_p^w), \\ L \cdot \sigma_p^w &\leq c_p^w - \pi^w \leq U \cdot \sigma_p^w, \\ c_p^w - \pi^w &\geq 0, \\ \sigma_p^w &\in \{0, 1\}, \\ \forall p &\in R_w, \quad w \in W, \end{aligned} \quad (10)$$

where L is a large enough negative number; U is a large enough positive number; ε is a very small positive number; and σ_p^w is a 0-1 variable. If $\sigma_p^w = 0$, then $f_p^w > 0$; on the contrary, if $\sigma_p^w = 1$, f_p^w equals zero. The detailed verification of the equivalence between the set of constraints (10) and UE condition can be referred to the work of Lo [47, 48] and Wang and Lo [40]. Note that constraints (10) are now linear in f_p^w , σ_p^w , c_p^w , and π^w .

In the bilevel formulation, queuing delay d_a is implicit as mentioned above. To transform the bilevel program into single-level problem, d_a should be converted into an explicit variable. The following process shows how d_a is converted and how to formulate the mixed-integer constraints associated with (7).

Since queues may only occur on saturated roads, the following relationship should be enforced [22]:

$$\begin{aligned} \lambda_a S_a &\geq x_a \implies \\ d_a &= 0, \\ \lambda_a S_a &= x_a \implies \\ d_a &\geq 0, \end{aligned} \quad (11)$$

$$\forall a \in A,$$

which means if a link is unsaturated, that is, $\lambda_a S_a \geq x_a$, then the queuing delay on this link is zero; otherwise, if a link is oversaturated, the queuing delay is greater than zero. Constraint (11) is similar to that in (8). Therefore, constraint (11) can also be rewritten into the form of complementary problem in

$$\begin{aligned} (\lambda_a S_a - x_a) \cdot d_a &= 0, \\ \lambda_a S_a - x_a &\geq 0, \end{aligned} \quad (12)$$

$$\forall a \in A.$$

Then following the transformation procedure as imposed on (9), the constraints in terms of mixed-integer linear constraints in (13) can be developed to replace (12):

$$\begin{aligned} L \cdot \gamma_a + \varepsilon &\leq \lambda_a S_a - x_a \leq U \cdot (1 - \gamma_a), \\ L \cdot \gamma_a &\leq d_a \leq U \cdot \gamma_a, \\ \lambda_a S_a - x_a &\geq 0, \\ \gamma_a &\in \{0, 1\}, \\ \forall a &\in A, \end{aligned} \quad (13)$$

where γ_a is a binary variable. For $\gamma_a = 0$, we have

$$\begin{aligned} \gamma_a &= 0 \\ \Rightarrow \begin{cases} \varepsilon \leq \lambda_a S_a - x_a \leq U \iff \lambda_a S_a - x_a > 0 \\ 0 \leq d_a \leq 0 \iff d_a = 0. \end{cases} \end{aligned} \quad (14)$$

On the other hand, if $\sigma_p^w = 1$, since $d_a \geq 0$, we have

$$\gamma_a = 1 \Rightarrow \begin{cases} L + \varepsilon \leq \lambda_a S_a - x_a \leq 0 \iff \lambda_a S_a - x_a \leq 0 \\ L \leq d_a \leq U \\ d_a \geq 0 \\ \lambda_a S_a - x_a \geq 0 \end{cases} \Rightarrow \begin{cases} \lambda_a S_a - x_a = 0 \\ \iff d_a \geq 0. \end{cases} \quad (15)$$

Thus, the set of constraints (13) is equivalent to the “if-then” conditions in (11). Note that d_a is an explicit variable in constraints (13), which are linear in x_a , d_a , and γ_a but remain nonlinear in S_a and λ_a .

In summary, the single-level MIP for CNDSSP can be formulated as follows:

$$\begin{aligned} \text{Minimize } & F(\lambda, S, x(\lambda, S), d(\lambda, S)) \\ &= \sum_{a \in A^0} \rho x_a \cdot (t_a(x_a, \lambda_a, S_a) + d_a) \\ &+ \sum_{a \in A^0} g_a(S_a) \end{aligned} \quad (16)$$

subject to the following:

(i) Signal variable constraints:

$$\begin{aligned} \lambda_a^{\min} &\leq \lambda_a \leq \lambda_a^{\max}, \quad a \in A, \\ \sum_{a \in A_i} \lambda_a &= 1, \quad i \in I. \end{aligned} \quad (17)$$

(ii) Capacity constraints and demand conservation constraints:

$$\begin{aligned} S_a^0 &\leq S_a \leq S_a^{\max}, \quad a \in A, \\ d_a \cdot x_a &\leq Q_a(S_a), \quad \forall a \in A, \\ q^w &= \sum_{p \in R_w} f_p^w, \quad \forall w \in W. \end{aligned} \quad (18)$$

(iii) UE constraints:

$$\begin{aligned} L \cdot \sigma_p^w + \varepsilon &\leq f_p^w \leq U \cdot (1 - \sigma_p^w), \\ L \cdot \sigma_p^w &\leq c_p^w - \pi^w \leq U \cdot \sigma_p^w, \\ c_p^w - \pi^w &\geq 0, \\ \sigma_p^w &\in \{0, 1\}, \\ \forall p &\in R_w, \quad w \in W. \end{aligned} \quad (19)$$

(iv) Queuing delay constraints:

$$\begin{aligned} L \cdot \gamma_a + \varepsilon &\leq \lambda_a S_a - x_a \leq U \cdot (1 - \gamma_a), \\ L \cdot \gamma_a &\leq d_a \leq U \cdot \gamma_a, \\ \lambda_a S_a - x_a &\geq 0, \\ \gamma_a &\in \{0, 1\}, \\ \forall a &\in A. \end{aligned} \quad (20)$$

(v) Definitional constraints:

$$\begin{aligned} x_a &= \sum_{w \in W} \sum_{p \in R_w} \delta_{ap}^w \cdot f_p^w, \quad \forall a \in A^0, \\ c_p^w &= \sum_{a \in A} \delta_{ap}^w \cdot (t_a(x_a, \lambda_a, S_a) + d_a), \\ &\forall p \in R_w, \quad w \in W, \end{aligned} \quad (21)$$

$$\begin{aligned} x_a &\geq 0, \quad \forall a \in A^0, \\ d_a &\geq 0, \quad \forall a \in A, \\ f_p^w &\geq 0, \quad \forall p \in R_w, \quad w \in W. \end{aligned}$$

In the above single-level mixed-integer model for CNDSSP, the UE conditions and queuing delay constraints are represented by a set of mixed-integer constraints (19)-(20), rather than in a lower-level program formulation. One can note that the nonlinearity stems from the link travel time function, the objective function, and capacity and queuing delay constraints; that is, this single-level model is a mixed-integer nonlinear program (MINLP). In the next section, we will employ linearization technique to transform the MINLP into MILP, through which global optimal solution can be obtained.

3. Linearization of the Single-Level Model

Mixed-integer models would be used to formulate the nonlinear terms into piecewise-linear approximations. Indeed, many mixed-integer models have been studied to capture piecewise-linear functions [49–52]. However, different formulations have different properties and their computational efficiencies vary from each other. Recently, several new and existing modeling efforts of piecewise-linear functions are reviewed in the work of Vielma et al. [53]. In this study, we apply the logarithmic formulation Log model (logarithmic branching convex combination model) to linearize the nonlinear terms, which significantly reduces the number of variables and constraints needed by introducing a binary in $\{0, 1\}^{\lceil \log_2 |P| \rceil}$ through an injective function $P \rightarrow \{0, 1\}^{\lceil \log_2 |P| \rceil}$ to identify an active polytope P in the linearization process. Specifically, the logarithmic relationship between the number of binary variables and the number of polytopes and constraints makes sure that the size of the linearization model will not be prohibitively large even when high approximation accuracy and refined partition scheme are required.

This section includes three parts. The details of linearizing the two-dimensional functions and MILP formulation are expressed in Section 3.1. The objective function and the storage capacity function are then transformed in Section 3.2. The MILP model for the CNDSSP is summarized in Section 3.3. Finally, we adopt the MILP formulation for the case of discrete link capacity improvement.

3.1. Linearization of Two-Dimensional Functions. In the single-level MIP proposed above, there are several nonlinear terms that exist (i.e., $d_a \cdot x_a$, $\lambda_a \cdot S_a$, and $t_a(x_a, \lambda_a, S_a)$) and all of them are multivariate functions with multiple variables. In this study, the nonlinear terms will be approximated by piecewise-linear functions. In this way, the MINLP can be transformed into the MILP.

For the last nonlinear term $t_a(x_a, \lambda_a, S_a)$, it is used to cover both running time and signal delay at link exit. The signal delay stems from the signal interruption of traffic and could be determined by approximated formula developed from local traffic conditions, such as Webster's two-term delay formula for signalized intersections [54] and the polynomial cost function [17]. However, as was pointed out in Yang and Yagar [22], both of the two formulas have drawbacks in describing the signal delay, especially for the saturated links. Therefore, following Yang and Yagar [22], we use the link performance function to depict both the running time

on the link and signal delay at link exit. The formulation is given as follows:

For signal-controlled link $a \in A$:

$$t_a(x_a, \lambda_a, S_a) = t_a^0 \left[1 + \left(\frac{x_a}{\lambda_a S_a} \right)^2 \right]. \quad (22)$$

For non-signal-controlled link $a \in \bar{A}$:

$$t_a(x_a, S_a) = t_a^0 \left[1 + \left(\frac{x_a}{S_a} \right)^2 \right], \quad (23)$$

where t_a^0 is free-flow travel time of link a and given as known constants. Note that (22) is a function with three variables and it includes the term $\lambda_a \cdot S_a$ in its formulation. If the term $\lambda_a \cdot S_a$ is linearized first and substituted by a new variable, (22) will have the same form as (23) and both of them are two-dimensional. Since the other two nonlinear terms, that is, $\lambda_a \cdot S_a$ and $d_a \cdot x_a$, are both two-dimensional, all the nonlinear terms in the single-level model can be regarded as two-dimensional and be linearized through the same piecewise-linear approximation method. Now $\lambda_a \cdot S_a$ is taken as the example to demonstrate the detailed piecewise-linearization process by applying Log model.

Let $\lambda_a \cdot S_a$ be represented by a new variable B_a ; that is, $B_a = B(S_a, \lambda_a) = \lambda_a \cdot S_a$. The bounded domain of B_a would be partitioned into a finite family of polytopes P , which is usually taken to be a triangulation of the field [51]. For Log model, the family of polytopes needs to be topologically compatible or equivalent with a triangulation known as J_1 or "Union Jack" [55]. The Log model needs only logarithmic number of binary variables and thus a branching scheme with a logarithmic number of dichotomies, which is introduced in Vielma and Nemhauser [56]. In this paper, the domain of B_a is to be divided into subsets as shown in Figure 1 when $P = J_1$. The feasible region of S_a (i.e., $[S_a^0, S_a^{\max}]$) is partitioned into $2N$ intervals in x dimension; that is, $[S_a^{n-1}, S_a^n]$, $n = 1, 2, \dots, 2N$, and λ_a (i.e., $[\lambda_a^{\min}, \lambda_a^{\max}]$) is partitioned into $2M$ intervals in y dimension; that is, $[\lambda_a^{m-1}, \lambda_a^m]$, $m = 1, 2, \dots, 2M$. Note that the lengths of these intervals are not necessarily the same.

Apparently, with the increasing number of N and M , the piecewise-linear approximation will be more accurate. However, not only does the approximation accuracy depend on the total number of polytopes; the detailed partition scheme also impacts the results remarkably. For example, in Figure 2, although the piecewise-linear function L_1 only has two pieces, the minimum of L_1 is equal to the actual minimum of original function. As compared to L_1 , the minimum of L_2 , which has three segments, is larger than the actual minimum. Hence, advanced partition method, which better describes the profile characteristics of the graph of nonlinear functions, can be studied in the future. In this paper, we just apply the commonly used evenly partitioned scheme.

FIGURE 1: An equivalent J_1 triangulation of the feasible region. FIGURE 2: Different piecewise-linear approximations for $f(x)$.

Following Log method in Vielma et al. [53], the function $B_a = B(S_a, \lambda_a) = \lambda_a \cdot S_a$ for $a \in A$ can be substituted by a set of MILP constraints as follows:

$$\sum_{v \in V(P)} \alpha_v (v_x, v_y) = (S_a, \lambda_a), \quad (24)$$

$$\sum_{v \in V(P)} \alpha_v B(v_x, v_y) = B_a,$$

$$\alpha_v \geq 0 \quad \forall v \in V(P),$$

$$\sum_{v \in V(P)} \alpha_v = 1, \quad (25)$$

$$\begin{aligned} \sum_{v \in L_k} \alpha_v &\leq y_k, \\ \sum_{v \in R_k} \alpha_v &\leq 1 - y_k, \\ y_k &\in \{0, 1\} \quad \forall k \in K. \end{aligned} \quad (26)$$

In the above formulation, P is the family of partitioned triangles and $v \in V(P)$ denotes corner-point, which may belong to more than one triangle in the partitioned domain; (v_x, v_y) represents the coordinates of corner-point v ; α_v is the weighted factor of convex combination related to corner-point v , and, hence, $\alpha_v \in [0, 1]$. Since B_a must fall into a single triangle of the domain, that is, the active triangle, B_a can be approximated in the form of a weighted function of $B(v_x, v_y)$ value of vertices of this triangle; (S_a, λ_a) can also be calculated from a weighted function of the coordinates of the vertices of the active triangle, as is shown in (24). Equation (25) entails that the sum of weighted factors α_v of all vertices in the domain is equal to 1. Indeed, only α_v of the active triangle is larger than zero while the weighted factors of the other inactive triangles are all equal to zero.

As α_v is not associated with any specific polytope, constraints (24)-(25) only require the nonnegativity of α_v . However, constraints (26) ensure that all nonzero α_v are related to the corner-points of one polytope in P . In (26), $L_k, R_k \subset V(P)$ and the set of dichotomies $\{L_k, R_k\}_{k \in K}$ is referred to a branching scheme for the family of polytopes P that is equivalent to J_1 . For each particular triangle $P^* \in P$, the vertices of P^* can be expressed by $V(P^*) = \bigcap_{k \in K} (V(P) \setminus T_k)$, where $T_k = L_k$ or $T_k = R_k$ for each $k \in K$. The index set K is divided into two subsets K_1 and K_2 . The first is given by $K_1 = \{(kx, ky) \mid kx \in \{1, \dots, \text{Dim}\}, ky \in \{1, \dots, \lceil \log_2 N_{kx} \rceil\}\}$, where Dim refers to dimension number of the domain and it is equal to two in this case because our nonlinear function is two-dimensional; N_{kx} is the number of intervals divided in kx dimension and note that N_{kx} is even because of J_1 triangulation. Assume that each interval index $n \in \{1, 2, \dots, N_{kx}\}$ can be represented by a fixed binary vector $G^n \in \{0, 1\}^{\lceil \log_2 N_{kx} \rceil}$. The selection of G^n is arbitrary on condition that the two successive vectors G^n and G^{n+1} differ in only one bit for each $n \in \{1, 2, \dots, N_{kx} - 1\}$. The set of binary vectors with this property is known as Gray code or reflected binary code [57]. Then, for each $(kx, ky) \in K_1$, $L_{(kx, ky)} = \{v \in V \mid v_{kx} \in O(ky, 1)\}$ and $R_{(kx, ky)} = \{v \in V \mid v_{kx} \in O(ky, 0)\}$, where $O(l, b) = \{n \in \{0, 1, \dots, N_{kx}\} \mid (n = 0 \text{ or } G_l^n = b) \text{ and } (n = N_{kx} \text{ or } G_l^{n+1} = b)\}$. The second set K_2 is denoted by $K_2 = \{(kx, ky) \mid kx, ky \in \{1, \dots, \text{Dim}\} \text{ and } kx < ky\}$. For each $(kx, ky) \in K_2$, $L_{(kx, ky)} = \{v \in V \mid v_{kx} \text{ is even and } v_{ky} \text{ is odd}\}$ and $R_{(kx, ky)} = \{v \in V \mid v_{kx} \text{ is odd and } v_{ky} \text{ is even}\}$. The size of (26) is equal to $2[\text{Dim} \cdot \lceil \log_2 N_{kx} \rceil + \text{Dim}(\text{Dim} - 1)/2]$ and only dependent on $\lceil \log_2 N_{kx} \rceil$ when Dim is fixed. A simple example of this branching scheme of an equivalent J_1 triangulation is given in detail in the appendix.

Similarly, a new variable H_a is introduced to denote the term of $d_a \cdot x_a$ in the single-level problem. Then, we apply Log model for piecewise-linear approximation, and the function

$H_a = H(d_a, x_a) = d_a \cdot x_a$ for $a \in A$ can be written into a set of MILP constraints as follows:

$$\begin{aligned}
 \sum_{v \in V(P_1)} \beta_v (v_x, v_y) &= (d_a, x_a), \\
 \sum_{v \in V(P_1)} \beta_v H(v_x, v_y) &= H_a, \\
 \beta_v &\geq 0 \quad \forall v \in V(P_1), \\
 \sum_{v \in V(P_1)} \beta_v &= 1, \\
 \sum_{v \in L_k} \beta_v &\leq z_k, \\
 \sum_{v \in R_k} \beta_v &\leq 1 - z_k, \\
 z_k &\in \{0, 1\} \quad \forall k \in K,
 \end{aligned} \tag{27}$$

where $P_1 = J_1$ is the family of polytopes, which are partitioned triangles of the domain of H_a ; that is, $D_H = \{(d_a, x_a) \mid d_a \in [d_a^{\min}, d_a^{\max}], x_a \in [x_a^{\min}, x_a^{\max}]\}$. For link flow x_a , the upper bound $x_a^{\max}$ can be set as the sum of travel demands of the OD pairs, at least one viable path of which includes link a , or simply the sum of all OD demands. Since OD demands are given as constants in our model, the upper bound $x_a^{\max}$ can be easily defined. As for queuing delay d_a , there is no explicit upper bound in the constraints. However, considering the realistic condition of queuing delay at an intersection, d_a should not be unbounded. $d_a^{\max}$ can be set as a positive large enough number which fully covers the feasible region of d_a , for example, four times signal cycle or free-flow travel time of the whole link a . Both $x_a^{\min}$ and $d_a^{\min}$ are set equal to zero.

Now consider the linearization of link travel time function. For signal-controlled link, $\lambda_a \cdot S_a$ is substituted by B_a . Thus, the link travel time function can be regarded as a function with respect to both x_a and B_a , that is, $t_a(x_a, B_a) = t_a^0[1 + (x_a/B_a)^2]$, $a \in A^0$, the piecewise-linear approximation of which is represented as follows:

$$\begin{aligned}
 \sum_{v \in V(P_2)} \gamma_v (v_x, v_y) &= (x_a, B_a), \\
 \sum_{v \in V(P_2)} \gamma_v t_a(v_x, v_y) &= t_a, \\
 \gamma_v &\geq 0 \quad \forall v \in V(P_2), \\
 \sum_{v \in V(P_2)} \gamma_v &= 1, \\
 \sum_{v \in L_k} \gamma_v &\leq u_k, \\
 \sum_{v \in R_k} \gamma_v &\leq 1 - u_k, \\
 u_k &\in \{0, 1\} \quad \forall k \in K,
 \end{aligned} \tag{28}$$

where $P_2 = J_1$ is the family of polytopes which is the triangulation of t_a . The range of B_a is defined as $[B_a^{\min}, B_a^{\max}]$, where $B_a^{\min} = \lambda_a^{\min} \cdot S_a^0$ and $B_a^{\max} = \lambda_a^{\max} \cdot S_a^{\max}$. For non-signal-controlled link $a \in \bar{A}$, since $t_a(x_a, S_a) = t_a^0[1 + (x_a/S_a)^2]$ is originally two-dimensional, the piecewise-linear functions can still be expressed as (28) by changing B_a to S_a .

3.2. Linearization of Objective Function and Storage Capacity Function. The objective function of the single-level optimization model is the combination of actual total system travel cost and associated link expansion cost. The first term of the objective, that is, $\sum_{a \in A^0} \rho x_a \cdot (t_a(x_a, \lambda_a, S_a) + d_a)$, is nonlinear and nonconvex. But, through the UE principle, it is easy to rewrite the first part into a linear form. When UE condition is satisfied, all used paths of the same OD pair must be equal to the minimum travel time π^w , which includes running time, signal delay, and queuing delay in this model. Then, the total system travel cost can be rewritten into the equivalent form of $\sum_{w \in W} \rho q^w \pi^w$, where all OD demands q^w are given constants and ρ is the factor to convert travel time to monetary travel cost, which is also a constant. Thus, the first term of the objective function is linear. The second term of the objective function $\sum_{a \in A^0} g_a(S_a)$, which determines the link expansion expense, is a function of S_a . In practical use, due to various topographic and geological conditions of land, the link investment cost function can be quite different. If $g_a(S_a)$ is linear, the whole objective function is linear, and no transformation is needed. Otherwise, the linearization method Log applied in the last section can be used to transform the nonlinear functions into piecewise-linear approximations. As $g_a(S_a)$ is only a single variable function, the linearization process is much easier than multivariate functions, like $t_a(x_a, \lambda_a, S_a)$. The domain of g_a is one-dimensional and partitioned into line segments rather than triangles. In conclusion, the investment function can also be expressed in a linear form and thus the whole objective function is linear.

For the link storage capacity function $Q_a(S_a)$, which is also a single variable function with respect to S_a , it can be treated in the identical way as in the link expansion cost function. That is, we can always have a linear or piecewise-linearized link storage capacity function Q_a . In this paper, for illustration purpose, we simply assume linear functions of $g_a(S_a)$ and $Q_a(S_a)$.

3.3. The Reformulated MILP Model. In summary, the reformulated MILP model of the original single-level optimization program is summarized as follows:

$$\text{Minimize } F(\lambda, S, x, d) = \sum_{w \in W} \rho q^w \pi^w + \sum_{a \in A^0} g_a(S_a) \tag{29}$$

subject to the following:

- (i) Green split constraints for links with signal control are as follows:

$$\begin{aligned}
 \lambda_a^{\min} &\leq \lambda_a \leq \lambda_a^{\max}, \quad a \in A, \\
 \sum_{a \in A_i} \lambda_a &= 1, \quad i \in I.
 \end{aligned} \tag{30}$$

- (ii) Capacity constraints and demand conservation constraints are as follows:

$$\begin{aligned} S_a^0 &\leq S_a \leq S_a^{\max}, \quad a \in A^0, \\ H_a &\leq Q_a(S_a), \quad \forall a \in A, \\ q^w &= \sum_{p \in R_w} f_p^w, \quad \forall w \in W. \end{aligned} \quad (31)$$

- (iii) UE condition constraints are as follows:

$$\begin{aligned} L \cdot \sigma_p^w + \varepsilon &\leq f_p^w \leq U \cdot (1 - \sigma_p^w), \\ L \cdot \sigma_p^w &\leq c_p^w - \pi^w \leq U \cdot \sigma_p^w, \\ c_p^w - \pi^w &\geq 0, \\ \sigma_p^w &\in \{0, 1\}, \\ \forall p &\in R_w, \quad w \in W. \end{aligned} \quad (32)$$

- (iv) Queuing delay constraint is as follows:

$$\begin{aligned} L \cdot \gamma_a + \varepsilon &\leq B_a - x_a \leq U \cdot (1 - \gamma_a), \\ L \cdot \gamma_a &\leq d_a \leq U \cdot \gamma_a, \\ x_a &\leq B_a, \\ \gamma_a &\in \{0, 1\}, \\ \forall a &\in A. \end{aligned} \quad (33)$$

- (v) Linearization of function $B_a = \lambda_a S_a$ for $a \in A$ is as follows:

$$\begin{aligned} \sum_{v \in V(P)} \alpha_v (v_x, v_y) &= (S_a, \lambda_a), \\ \sum_{v \in V(P)} \alpha_v B(v_x, v_y) &= B_a, \\ \alpha_v &\geq 0 \quad \forall v \in V(P), \\ \sum_{v \in V(P)} \alpha_v &= 1, \\ \sum_{v \in L_k} \alpha_v &\leq \gamma_k, \\ \sum_{v \in R_k} \alpha_v &\leq 1 - \gamma_k, \\ \gamma_k &\in \{0, 1\} \quad \forall k \in K. \end{aligned} \quad (34)$$

- (vi) Linearization of function $H_a = d_a x_a$ for $a \in A$ is as follows:

$$\begin{aligned} \sum_{v \in V(P_1)} \beta_v (v_x, v_y) &= (d_a, x_a), \\ \sum_{v \in V(P_1)} \beta_v H(v_x, v_y) &= H_a, \\ \beta_v &\geq 0 \quad \forall v \in V(P_1), \end{aligned}$$

$$\begin{aligned} \sum_{v \in V(P_1)} \beta_v &= 1, \\ \sum_{v \in L_k} \beta_v &\leq z_k, \\ \sum_{v \in R_k} \beta_v &\leq 1 - z_k, \\ z_k &\in \{0, 1\} \quad \forall k \in K. \end{aligned} \quad (35)$$

- (vii) For links with signal control $a \in A$, linearization of the link travel time function t_a is as follows, while, for links without signal control $a \in \tilde{A}, B_a$ in the following constraints is substituted by S_a :

$$\begin{aligned} \sum_{v \in V(P_2)} \gamma_v (v_x, v_y) &= (x_a, B_a), \\ \sum_{v \in V(P_2)} \gamma_v t_a(v_x, v_y) &= t_a, \\ \gamma_v &\geq 0 \quad \forall v \in V(P_2), \\ \sum_{v \in V(P_2)} \gamma_v &= 1, \\ \sum_{v \in L_k} \gamma_v &\leq u_k, \\ \sum_{v \in R_k} \gamma_v &\leq 1 - u_k, \\ u_k &\in \{0, 1\} \quad \forall k \in K. \end{aligned} \quad (36)$$

- (viii) Demand conservation and definitional constraints are as follows:

$$\begin{aligned} x_a &= \sum_{w \in W} \sum_{p \in R_w} \delta_{ap}^w \cdot f_p^w, \quad \forall a \in A^0, \\ c_p^w &= \sum_{a \in A^0} \delta_{ap}^w \cdot (t_a(x_a, \lambda_a, S_a) + d_a), \\ \forall p &\in R_w, \quad w \in W, \\ x_a &\geq 0, \quad \forall a \in A^0, \\ d_a &\geq 0, \quad \forall a \in A, \\ f_p^w &\geq 0, \quad \forall p \in R_w, \quad w \in W. \end{aligned} \quad (37)$$

So far, the CNDSSP model with queuing in saturated network is cast into mixed MILP. It can thus be easily solved by existing commercial MILP solvers like CPLEX, and the solution possesses global optimization property. As the MILP is an approximation of the original model by using a linearization approach, this obtained solution can also be treated as the approximated global optimal solution of the original problem. In the linearization and approximation process, an error is inevitably introduced, which can be reduced by

adopting a finer linearization scheme. Therefore, the global optimality of the solution determined with this approach is subject to the linearization scheme adopted, which can be fully prespecified according to the desired accuracy requirement. In other words, the finer the linearization scheme adopted, the more refined the global optimal solution found. In this way, we contend that the proposed method is able to find the global optimum of the CNDSSP, up to the accuracy of the linearization scheme specified or adopted.

3.4. Adapting the MILP for Discrete Link Capacity Improvement. In reality, the expansion of roadway segments may be on discrete levels, which means that the road capacity expansion is determined by the number of lanes added. To deal with the CNDSSP on discrete levels of link capacity improvement, we will make some amendments on the above MILP model to adapt it for solving this problem.

With discrete link capacity improvement, the CNDSSP changes to decide how many lanes should be added on each link in addition to setting signal timings, which is a mixed discrete and continuous problem. Let Ln_a denote the number of lanes to be added to link $a \in A^0$ and let LN_a be the maximum number of lanes that can be added on link a . Supposing the capacity raised by each additional lane on link a is represented by a constant Lc_a , thus the discrete variable of link capacity can be written as

$$S_a = S_a^0 + Ln_a \cdot Lc_a, \quad \forall a \in A^0 \quad (38)$$

wherein S_a^0 is the original capacity of link a , $Ln_a \in \{1, 2, \dots, LN_a\}$. Hence, the restriction of road capacity is decided by the maximum number of lanes LN_a added to that link and then the first constraint of (31) can be removed. Now, we only need to further express the integer variable Ln_a into a combination of a series of binary variables, as is shown below:

$$\begin{aligned} Ln_a &= \sum_{i=1}^{LN_a} i \cdot \zeta_a^i, \\ \sum_{i=1}^{LN_a} \zeta_a^i &= 1, \\ \zeta_a^i &\in \{0, 1\}, \quad \forall i \in \{1, 2, \dots, LN_a\}, \\ \forall a &\in A^0, \end{aligned} \quad (39)$$

where ζ_a^i is the binary variable that represents adding a number of i lanes on link a . Constraints (39) state that if a number of j lanes are added, that is, $Ln_a = j$, then $\zeta_a^j = 1$ and all the other binary variables $\zeta_a^i, \forall i = 1, 2, \dots, LN_a$ and $i \neq j$, are equal to zero. Thus, the number of additional lanes Ln_a is determined by the set of ζ_a^i . As constraints (38)–(39) are linear, they can be directly attached to MILP (29)–(37) and then be solved by commercial software. In conclusion, through the adaptation made above, the MILP for CNDSSP with continuous road capacity improvement can also be extended to solve the CNDSSP with discrete link capacity.

TABLE 1: Link data input for the small network.

Link	Free-flow time	Expansion cost factor	Current capacity	Bound of capacity expansions
1	4	2	60	80
2	5	3	50	80
3	2	5	50	70
4	8	4	50	60
5	8	1	80	100
6	4	3	60	90
7	4	6	50	70
8	2	2	60	80
9	3	4	50	70

TABLE 2: Other parameters used in the network.

Variable	Value
Demand	
1 $\rightarrow$ 5	70
2 $\rightarrow$ 4	80
Bound of green split	
$\lambda_a^{\min}$	0.05
$\lambda_a^{\max}$	0.95
Maximum storage capacity multiplier:	
$Q_a(S_a) = k_a \cdot S_a$	
k_1	1.7
k_2	2.0

4. Numerical Studies

In the numerical studies, the performance of the proposed CNDSSP model and the Log-MILP method is tested by three examples in different networks. We also compare the results of other models and local optimal method. The Log-MILP solution method of the three cases is modeled by YALMIP-R20120420 [58] and then solved by external commercial optimization solver CPLEX optimization studio 12.3 [59]. All experiments are undertaken on a personal computer with Intel Core i7 860 at 2.80 GHz CPU, 8 GB RAM, and Windows 7 Enterprise operating system (64-bit).

4.1. Case 1: The Yang-Yagar Network. This case study follows the numerical example used in Yang and Yagar [22]. Through this example, we compare our CNDSSP model with two other models: ENSS and CNDP. We also compare the global optimal solution from our Log-MILP method with result from the method of sensitivity analysis, which is indeed local optimal method. The small test network with one signal-controlled intersection (node 3) is depicted in Figure 3. Link parameters including free-flow travel time, expansion cost factor, current capacity, and expansion bound are listed in Table 1. OD demands, bound of green splits, and link storage capacity multiplier are all shown in Table 2.

4.1.1. Comparison of CNDSSP Model with Other Models. In solving the proposed CNDSSP model by using the Log-MILP

TABLE 3: Comparison of CNDDSP with other models.

Variable	Do nothing	ENSS	CNDP	CNDDSP
Green split of link 1	0.5	0.2333	0.5	0.3985
Capacity enhancement				
Δs_1	—	—	13.2500	13.9359
Δs_2	—	—	26.2500	24.0962
Δs_5	—	—	20	20
Δs_8	—	—	0.0873	
Queuing delay				
d_1	1.7439	3.3883		
d_2	2.3188	0.6030		
Objective value				
F	—	2398.8	2177.1	2168.8
F_r	2415.2	2398.7	2177.0	2168.6

FIGURE 3: The small Yang-Yagar network.

algorithm, since feasible regions of x_a , s_a , and B_a are much larger than those of λ_a and d_a , the number of partitioned segments of x_a , s_a , and B_a should also be larger to ensure the accuracy as well as the computation speed at the same time. Here, the domains for x_a , s_a , and B_a are all separated into 80 equal segments, whereas those of λ_a and d_a are both separated into 20 equal segments. This partition scheme is adopted throughout the tests with the solution method of Log-MILP in Section 4.1. Partition scheme has great effects on the accuracy of solutions and the computational time. If a finer solution is needed, a much more refined partition scheme may be selected. It is obvious that the bounds of the variables should be tight to avoid excessive branching and thus increased computational time. The upper bound of x_a is set as the sum of the demands of OD pairs, at least one path of which involves the link. Assume the upper bound of d_a is 10 which is even larger than the maximum link free-flow travel time.

Table 3 displays the optimal solution and objective value of CNDDSP solved by the proposed Log-MILP method. The optimal solution involves link improvement and signal settings (suppose the original green ratio of link 1 is 0.5) and F is the objective value calculated directly from the Log-MILP. Table 3 shows the exact objective function values recalculated from the UE traffic assignment with side constrained link capacity [60] by substituting the optimal solution of link capacity and signal setting from the Log-MILP, which are

indicated by F_r . In order to compare the proposed CNDDSP model with the traditional CNDP and the ENSS model, the test network is improved by applying the three models individually. All these models are solved by the Log-MILP algorithm. Table 3 also reports the final results obtained from these models. Before any model is applied to the test network, that is, the case labeled “do nothing” in Table 3, the value of the objective function is 2415.2 and the queuing delay on link 1 and link 2 is 1.7439 and 2.3188, respectively [22]. When the ENSS model is imposed, the green split of link 1 is shifted to 0.2333, which reduces the objective value to 2398.8. This model improves the performance of the network by 0.68% ($= [2398.8 - 2415.2]/2415.2$). However, the queuing delay on both two approaching legs (i.e., link 1 and link 2) still exists and the delay on link 1 even increases. When using the CNDP model, the optimal solution involves the capacity enhancement on links 1, 2, 5, and 8, supposing the green ratio of link 1 is fixed. The value of objective function decreases greatly to 2177.1 and the performance of the network increases by 9.86% ($= [2177.1 - 2415.2]/2415.2$). Besides, the queuing delay on both link 1 and link 2 reduces to zero. When the proposed CNDDSP model is applied to the network, the objective value is the best among the results of the three models, which decreases by 10.20% ($= [2168.8 - 2415.2]/2415.2$) from the do-nothing case, and there is also no queuing delay on both two links. The gaps between the objective values directly from the proposed Log-MILP method and the exact ones are no larger than 0.0092% (from the CNDDSP model, $= [2168.8 - 2168.6]/2168.6$), which are quite small. This may be due to the small size of the test network. The computational time for the three models is 0.21 min (ENSS), 1.46 min (CNDP), and 7.75 min (CNDDSP), respectively.

4.1.2. Comparison of Solution Results from Log-MILP Method with Sensitivity Analysis Approach. It is hard to test the performance of the Log-MILP method through the CNDDSP model because there is no benchmark algorithm to solve this problem in the literature. However, the ENSS problem can be seen as a special case of CNDDSP, where the link capacity is fixed and only the green split is variable. Since the ENSS problem has been tested with sensitivity analysis solution approach in Yang and Yagar [22], it is adopted to be compared with the presented global optimization algorithm. The former is a typical local heuristic algorithm. It is efficient and widely implemented in solving bilevel optimization problem, though only local optimal points may be found due to nonlinear nonconvex character of the bilevel problem.

According to the literature, for optimal solution from sensitivity analysis, the green split of link 1 is 0.2361 and the objective value is 2400.0 [22] after the iteration process, while, for optimal solution from Log-MILP, the green split of link 1 is 0.2333 and the corresponding final objective value is 2398.8. Obviously, Log-MILP yields a better objective value than that from sensitivity analysis approach, which is reduced by 0.0500% ($= [2398.8 - 2400.0]/2400.0$) compared to the result of the former method. Besides, the computational process only needs 0.21 min, which is very fast for solving this nonconvex problem.

TABLE 4: Results of the CNDSSP with discrete capacity enhancement applying the Log-MILP method.

Variable	Value
Green split of link 1	0.4043
Capacity enhancement	
Δs_1	20
Δs_2	20
Δs_5	20
Objective value	
F	2169.8
F_r	2169.6

4.1.3. An Example of Discrete Link Capacity Improvement. We show the results of the CNDSSP with discrete link capacity enhancement on the same test network in Figure 3. Suppose that each lane added possesses the capacity of 10 units and at most two more lanes can be added to each link. The proposed Log-MILP method is applied to solve the problem and the partition scheme remains the same as is used in the above tests. The results of optimal discrete capacity improvements are exhibited in Table 4. The optimal solution of this case study involves the addition of two more lanes on each of links 1, 2, and 5. The other roadway segments will not be expanded. The green split of link 1 is changed to 0.4043. This optimal solution results in no queuing delay on any signal-controlled link of the network. The value of the objective function achieves 2169.8, which is slightly higher than the result of the CNDSSP model with continuous capacity enhancement. This is because the discrete levels of capacity improvement impose tighter constraints on the structure of the network and thus lead to the objective value being not as good as that of the continuous CNDSSP model. However, the adjustment of signal setting can offset part of this effect and make the objective value better. The gap between the objective function value from Log-MILP method and the exact objective value is only 0.0092% ($= [2169.8 - 2169.6]/2169.6$), which is equal to the gap in the continuous capacity enhancement situation. The computational time of this example is 4.39 min.

4.2. Case 2: The Nguyen-Dupuis Network. To further illustrate the performance of the presented CNDSSP model and the Log-MILP solution method, the second case is applied to a larger-scale network. It is the Nguyen-Dupuis network, as shown in Figure 4, which is the same as was used in Nguyen and Dupuis [61]. The network has total four OD pairs, 19 links, and 25 paths. There are three signal-controlled intersections 6, 9, and 11 that are marked in the network and six signal-controlled links labeled from 1 to 6. The link parameters including free-flow travel time and expansion cost factors are listed in Table 5. Table 6 shows the demands of all OD pairs and maximum storage capacity multipliers of links with signal control. Let all current link capacities be 40 and let the maximum capacity of each link after expansions be 60. For the situation with discrete levels of capacity enhancements, it is assumed that each lane possesses ten units of link capacity and maximum two more lanes can be added to each

TABLE 5: Link parameters for the Nguyen-Dupuis network.

Link	Free-flow time	Expansion cost factor
1	5	3
2	2	2
3	2	3
4	3	3
5	2	2
6	3	4
7	1	6
8	2	2
9	1	4
10	2	6
11	3	2
12	2	2
13	1	4
14	1	4
15	4	1
16	2	3
17	1	7
18	5	1
19	2	5

TABLE 6: Demand and maximum storage capacities of signal-controlled links of the Nguyen-Dupuis network.

Maximum storage capacity multiplier: $Q_a(S_a) = k_a \cdot S_a$	Value	Demand	Value
k_1	2.3	1 → 2	34
k_2	1.7	4 → 2	46
k_3	1.8	1 → 3	44
k_4	2.0	4 → 3	26
k_5	2.2		
k_6	1.9		

link. Let the minimum green split for each signal-controlled link be 0.1 and let the maximum be 0.9. The conversion factor from travel time to money is set to be one and the upper bound of queuing delay of each link with signal is supposed to be ten.

In this example, we discretize the three variables x_a , s_a , and B_a into 40 equal segments and λ_a and d_a into twelve equal increments. Table 7 illustrates the results solved from the Log-MILP method for the two scenarios with continuous and discrete capacity enhancements. For the case of continuous link capacity improvements, the direct objective function value and the exact one are 3003.4 and 3101.4, respectively. The gap between them is 3.16% ($= [3003.4 - 3101.4]/3101.4$). This gap is larger than that in Yang-Yagar example, which may be due to the coarse resolution of variable partition scheme. For the discrete levels of link capacity improvements, the objective function value directly from the Log-MILP is 3061.5 and the exact objective value is 3116.7. The gap (1.77%) is smaller than the continuous scenario. In addition,

TABLE 7: Results generated from the Log-MILP method with continuous and discrete capacity enhancements.

Variable	Continuous capacity enhancement	Discrete capacity enhancement
Green split		
λ_1	0.5874	0.4563
λ_3	0.3207	0.1896
λ_5	0.6311	0.6229
Capacity enhancement		
Δs_1	20	20
Δs_2	20	20
Δs_3	20	20
Δs_4	20	20
Δs_5	19.9834	20
Δs_6		20
Δs_8	20	20
Δs_{10}	15.4667	20
Δs_{13}	9.2310	10
Δs_{15}	20	20
Δs_{18}	20	20
Queuing delay		
d_3	0.3802	3.0119
d_6	4.5345	2.7815
Objective value		
F	3003.4	3061.5
F_r	3101.4	3116.7

FIGURE 4: The Nguyen-Dupuis network.

the computational times for solving the continuous and the discrete case are 10.41 h and 12.30 h, respectively.

4.3. Case 3: The Sioux-Falls Network. To test the performance of the presented CNDSSP model and the Log-MILP solution method on large-size network, the widely used Sioux-Falls network is applied, which is shown in Figure 5. The network has total 528 nonzero OD pairs, 76 links, and 24 nodes. The value of network parameters is set basically following Suwansirikul et al. [6]. For non-signal-controlled links $a \in \bar{A}$, the link travel time function $t_a(x_a, S_a) = t_a^0[1 + 0.15(x_a/S_a)^4]$ is used, while, for signal-controlled links $a \in A$,

FIGURE 5: The Sioux-Falls network.

the formulation of link travel time function is $t_a(x_a, \lambda_a, S_a) = t_a^0[1 + 0.15(x_a/\lambda_a S_a)^4]$. Node 11 is set as a signal-controlled intersection and a simple two-phase signal system is adopted, which means links 27 and 36 share one phase (λ_1) and links 10 and 40 share the other (λ_2). Let the lower bound and upper bound of green split be 0.3 and 0.7 separately and let k_a for $a \in A$ in the maximum link storage capacity function be 1. The link parameters including original link capacity and free-flow travel time are listed in Table 8. Table 9 shows the demands of all 528 nonzero OD pairs. Assume link pairs (6, 8), (10, 16), and (13, 24) to be candidate links for capacity expansion. The construction cost function for candidate links and their separate upper bounds on link capacity after expansion are given in Table 10. The conversion factor from travel time to money is set to be one. In this example, we discretize x_a , s_a , and B_a into 10 equal segments and λ_a and d_a into four equal increments.

In this case, only the scenario with continuous capacity enhancement is tested. The result link capacity expansion and signal setting plan is shown in Table 11. The green split for phase 1 of the signal-controlled intersection should be set as 0.7, that is, on the upper boundary of the variable, which means reducing this value will increase the overall queuing delay and travel time or even violate the link exit capacity constraints. For candidate links to be expanded, capacities of both link pairs (6, 8) and (13, 24) are set as their maximum value, while the capacity of the pair (10, 16) is only planned to be enhanced by 4.2329. The objective value of the test directly from our solution approach is 94.7254, quite close to the real one, that is, 94.6954. The computational time of this test is about 16 hours, which is not very attractive for practical

TABLE 8: Link capacity for the Sioux-Falls network.

Link	Capacity	t_a^0
1	25.9002	0.06
2	23.4034	0.04
3	25.9002	0.06
4	4.9581	0.05
5	23.4034	0.04
6	17.1105	0.04
7	23.4034	0.04
8	17.1105	0.04
9	17.7827	0.02
10	9.8176	0.06
11	17.7827	0.02
12	4.9479	0.04
13	10.0000	0.05
14	4.9581	0.05
15	4.9479	0.04
16	4.8985	0.02
17	7.8418	0.03
18	23.4034	0.02
19	4.8985	0.02
20	7.8418	0.03
21	5.0501	0.10
22	5.0458	0.05
23	10.0000	0.05
24	5.0501	0.10
25	13.9157	0.03
26	13.9157	0.03
27	20.0000	0.05
28	13.5120	0.06
29	5.1334	0.05
30	4.9935	0.08
31	4.9088	0.06
32	10.0000	0.05
33	4.9088	0.06
34	4.8765	0.04
35	23.4034	0.04
36	9.8176	0.06
37	25.9002	0.03
38	25.9002	0.03
39	5.0912	0.04
40	9.7530	0.04
41	5.1275	0.05
42	4.9247	0.04
43	13.512	0.06
44	5.1275	0.05
45	15.6508	0.04
46	10.3149	0.04
47	5.0458	0.05
48	5.1334	0.05
49	5.2299	0.02
50	19.6798	0.03

TABLE 8: Continued.

Link	Capacity	t_a^0
51	4.9935	0.08
52	5.2299	0.02
53	4.8239	0.02
54	23.4034	0.02
55	19.6798	0.03
56	23.4034	0.04
57	15.6508	0.04
58	4.8239	0.02
59	5.0026	0.04
60	23.4034	0.04
61	5.0020	0.04
62	5.0599	0.06
63	5.0756	0.05
64	5.0599	0.06
65	5.2299	0.02
66	4.8853	0.03
67	10.3149	0.04
68	5.0756	0.05
69	5.2299	0.02
70	5.0000	0.04
71	4.9247	0.04
72	5.0000	0.04
73	5.0785	0.02
74	5.0912	0.04
75	4.8853	0.03
76	5.0785	0.02

use especially for a quite large network. Indeed, though the presented solution method can guarantee the global optimal solution, the solution efficiency is compromised and the significant improvement of the solution efficiency requires further study. The resulting link flow of the obtained optimal network plan is listed in Table 12.

5. Conclusions

This paper proposed an urban transport planning problem (i.e., CNDSSP) with combined network design and signal setting. Spillback phenomenon is considered in saturated conditions by controlling the queuing length on roadway segments to avoid blocking upstream links. A bilevel program model is developed to formulate this problem and then is transformed into an equivalent single-level MIP. In order to obtain a global optimal solution, an efficient mixed-integer linearization technique is applied to cast the problem into MILP formulation so that the solution possesses the property of global optimization. Finally, numerical experiments from the literature have been carried out and compared with the NDP and the ENSS models and a local optimal solution algorithm. Results show that all the tested cases can be solved efficiently via the commercial optimization software CPLEX. The optimal total system travel cost of CNDSSP formulation is less than those from pure NDP and pure ENSS model. In addition,

TABLE 9: Travel demands of the Sioux-Falls network.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	0	0.11	0.11	0.55	0.22	0.33	0.55	0.88	0.55	1.43	0.55	0.22	0.55	0.33	0.55	0.55	0.44	0.11	0.33	0.33	0.11	0.44	0.33	0.11
2	0.11	0	0.11	0.22	0.11	0.44	0.22	0.44	0.22	0.66	0.22	0.11	0.33	0.11	0.11	0.44	0.22	0	0.11	0.11	0	0.11	0	0
3	0.11	0.11	0	0.22	0.11	0.33	0.11	0.22	0.11	0.33	0.33	0.22	0.11	0.11	0.11	0.22	0.11	0	0	0	0	0.11	0.11	0
4	0.55	0.22	0.22	0	0.55	0.44	0.44	0.77	0.77	1.32	1.54	0.66	0.66	0.55	0.55	0.88	0.55	0.11	0.22	0.33	0.22	0.44	0.55	0.22
5	0.22	0.11	0.11	0.55	0	0.22	0.22	0.55	0.88	1.1	0.55	0.22	0.22	0.11	0.22	0.55	0.22	0	0.11	0.11	0.11	0.22	0.11	0
6	0.33	0.44	0.33	0.44	0.22	0	0.44	0.88	0.44	0.88	0.44	0.22	0.22	0.11	0.22	0.99	0.55	0.11	0.22	0.33	0.11	0.22	0.11	0.11
7	0.55	0.22	0.11	0.44	0.22	0.44	0	1.1	0.66	2.09	0.55	0.77	0.44	0.22	0.55	1.54	1.1	0.22	0.44	0.55	0.22	0.55	0.22	0.11
8	0.88	0.44	0.22	0.77	0.55	0.88	1.1	0	0.88	1.76	0.88	0.66	0.66	0.44	0.66	2.42	1.54	0.33	0.77	0.99	0.44	0.55	0.33	0.22
9	0.55	0.22	0.11	0.77	0.88	0.44	0.66	0.88	0	3.08	1.54	0.66	0.66	0.66	0.99	1.54	0.99	0.22	0.44	0.66	0.33	0.77	0.55	0.22
10	1.43	0.66	0.33	1.32	1.1	0.88	2.09	1.76	3.08	0	4.4	2.2	2.09	2.31	4.4	4.84	4.29	0.77	1.98	2.75	1.32	2.86	1.98	0.88
11	0.55	0.22	0.33	1.65	0.55	0.44	0.55	0.88	1.54	4.29	0	1.54	1.1	1.76	1.54	1.54	1.1	0.11	0.44	0.66	0.44	1.21	1.43	0.66
12	0.22	0.11	0.22	0.66	0.22	0.22	0.77	0.66	0.66	2.2	1.54	0	1.43	0.77	0.77	0.77	0.66	0.22	0.33	0.44	0.33	0.77	0.77	0.55
13	0.55	0.33	0.11	0.66	0.22	0.22	0.44	0.66	0.66	2.09	1.1	1.43	0	0.66	0.77	0.66	0.55	0.11	0.33	0.66	0.66	1.43	0.88	0.88
14	0.33	0.11	0.11	0.55	0.11	0.11	0.22	0.44	0.66	2.31	1.76	0.77	0.66	0	1.43	0.77	0.77	0.11	0.33	0.55	0.44	1.32	1.21	0.44
15	0.55	0.11	0.11	0.55	0.22	0.22	0.55	0.66	1.1	4.4	1.54	0.77	0.77	1.43	0	1.32	1.65	0.22	0.88	1.21	0.88	2.86	1.1	0.44
16	0.55	0.44	0.22	0.88	0.55	0.99	1.54	2.42	1.64	4.84	1.54	0.77	0.66	0.77	1.32	0	3.08	0.55	1.43	1.76	0.66	1.32	0.55	0.33
17	0.44	0.22	0.11	0.55	0.22	0.55	1.1	1.54	0.99	4.29	1.1	0.66	0.55	0.77	1.65	3.08	0	0.66	1.87	1.87	0.66	1.87	0.66	0.33
18	0.11	0	0	0.11	0	0.11	0.22	0.33	0.22	0.77	0.22	0.22	0.11	0.11	0.22	0.55	0.66	0	0.33	0.44	0.11	0.33	0.11	0
19	0.33	0.11	0	0.22	0.11	0.22	0.44	0.77	0.44	1.98	0.44	0.33	0.33	0.33	0.88	1.43	1.87	0.33	0	1.32	0.44	1.32	0.33	0.11
20	0.33	0.11	0	0.33	0.11	0.33	0.55	0.99	0.66	2.75	0.66	0.55	0.66	0.55	1.21	1.76	1.87	0.44	1.32	0	1.32	2.64	0.77	0.44
21	0.11	0	0	0.22	0.11	0.11	0.22	0.44	0.33	1.32	0.44	0.33	0.66	0.44	0.88	0.66	0.66	0.11	0.44	1.32	0	1.98	0.77	0.55
22	0.44	0.11	0.11	0.44	0.22	0.22	0.55	0.55	0.77	2.86	1.21	0.77	1.43	1.32	2.86	1.32	1.87	0.33	1.32	2.64	1.98	0	2.31	1.21
23	0.33	0	0.11	0.55	0.11	0.11	0.22	0.33	0.55	1.98	1.43	0.77	0.88	1.21	1.1	0.55	0.66	0.11	0.33	0.77	0.77	2.31	0	0.77
24	0.11	0	0	0.22	0	0.11	0.11	0.22	0.22	0.88	0.66	0.55	0.77	0.44	0.44	0.33	0.33	0	0.11	0.44	0.55	1.21	0.77	0

TABLE 10: Other candidate link parameters for the Sioux-Falls network.

Link pairs	(6, 8)	(10, 16)	(13, 24)	Construction cost function
$c_{\exp} (\times 10^{-3})$	26	48	34	$g_a(S_a) = c_{\exp}(S_a - S_a^0)^2$
$S_a^{\max}$	10.7242	12.9721	11.5024	

TABLE 11: Results of the Sioux-Falls network.

λ_1	0.70	d_{10}	0.25
λ_2	0.30	d_{40}	0.29
		d_{27}	0.24
		d_{36}	0.17
$\Delta s_{(6,8)}$	5.8257		
$\Delta s_{(10,16)}$	4.2329	F	94.7254
$\Delta s_{(13,24)}$	6.4112	F_r	94.6954

the proposed global optimal solution method obtains better urban network design and signal setting plan than a local optimal algorithm. However, it should be noted that most global optimal solution algorithms presented in the literature are not as efficient as local optimal solution methods and therefore not very appropriate for practical use. Hence, the improvement of the solution efficiency of the global optimal algorithm needs to be further addressed in the future studies.

FIGURE 6: A triangulation of the domain compatible with J_1 .

Appendix

A Simple Example of Branching Scheme for Equivalent J_1 Triangulation

Considering a nonlinear two-dimensional function $z = f(x, y)$, $x \in [0, 8]$, $y \in [1, 3]$, that needs to be linearized applying Log model, the entire domain of (x, y) is partitioned into 4×2 triangles as shown in Figure 6.

TABLE 12: Link flow result of the Sioux-Falls network.

Link	Flow
1	6.7518
2	9.3186
3	6.8982
4	7.8566
5	9.1722
6	19.3286
7	20.6234
8	20.3224
9	24.0234
10	2.9453
11	24.3617
12	9.3798
13	16.7351
14	8.0030
15	9.4043
16	19.6566
17	14.2038
18	17.3841
19	19.8276
20	13.9717
21	6.3608
22	9.1454
23	17.0488
24	6.2122
25	26.2220
26	26.5970
27	14.0019
28	22.4777
29	15.6947
30	8.4228
31	3.7108
32	14.0540
33	4.9250
34	3.9446
35	19.4832
36	6.8712
37	20.3146
38	21.2306
39	18.8525
40	2.9261
41	10.1642
42	9.2400
43	22.9063
44	9.8042
45	17.6190
46	18.7836
47	9.2329
48	15.9066
49	12.1426
50	19.9511

TABLE 12: Continued.

Link	Flow
51	7.9951
52	12.3777
53	10.3763
54	17.6162
55	19.9154
56	24.4013
57	17.7380
58	10.1838
59	9.3038
60	24.4878
61	9.2303
62	7.6960
63	8.3541
64	7.5996
65	10.1133
66	11.3879
67	18.6232
68	8.3534
69	10.2443
70	10.5437
71	8.5814
72	10.5137
73	10.2207
74	19.6585
75	11.1605
76	9.5321

For the first index set $(kx, ky) \in K_1$, when $kx = 1$, that is, kx refers to the x dimension, we have $ky \in \{1, \dots, \lceil \log_2 4 \rceil\} = \{1, 2\}$; when $kx = 1$, that is, kx refers to the y dimension, we have $ky \in \{1, \dots, \lceil \log_2 2 \rceil\} = \{1\}$. This implies that the Gray codes for intervals in x direction have two bits, whereas for intervals in y direction the Gray codes have only one bit. Assume $G^1 = (0, 0)$, $G^2 = (1, 0)$, $G^3 = (1, 1)$, and $G^4 = (0, 1)$ for $kx = 1$ and $G^1 = 0$ and $G^2 = 1$ for $kx = 2$.

When $kx = 1$ and $ky = 1$, for points with $x = 0$, since their index number $n = 0$ and $G_1^1 = 0$, then all of them belong to $R_{(1,1)}$; for points with $x = 2$, since $G_1^1 = 0$ and $G_1^2 = 1$, they belong to neither $L_{(1,1)}$ nor $R_{(1,1)}$; for points with $x = 4$, since $G_1^2 = 1$ and $G_1^3 = 1$, they all belong to $L_{(1,1)}$; for points with $x = 6$, since $G_1^3 = 1$ and $G_1^4 = 0$, they belong to neither $L_{(1,1)}$ nor $R_{(1,1)}$; for points with $x = 8$, since $G_1^4 = 0$ and $n = N_1 = 4$, they belong to $R_{(1,1)}$.

Therefore, for index set K_1 , we have

$$\begin{aligned} L_{(1,1)} &= \{(4, 1), (4, 2), (4, 3)\}, \\ R_{(1,1)} &= \{(0, 1), (0, 2), (0, 3), (8, 1), (8, 2), (8, 3)\}. \end{aligned} \quad (\text{A.1})$$

Similarly,

$$\begin{aligned} L_{(1,2)} &= \{(6, 1), (6, 2), (6, 3), (8, 1), (8, 2), (8, 3)\}, \\ R_{(1,2)} &= \{(0, 1), (0, 2), (0, 3), (2, 1), (2, 2), (2, 3)\}, \end{aligned}$$

$$\begin{aligned}
 L_{(2,1)} &= \{(0, 3), (2, 3), (4, 3), (6, 3), (8, 3)\}, \\
 R_{(2,1)} &= \{(0, 1), (2, 1), (4, 1), (6, 1), (8, 1)\}.
 \end{aligned}
 \tag{A.2}$$

For the second index set $(kx, ky) \in K_2$, since $kx, ky \in \{1, 2\}$ and $kx < ky$, we have $K_2 = \{(1, 2)\}$. $L_{(1,2)}$ is a set including those points with an even x index number and an odd y index number. $R_{(1,2)}$ is a set including points with an odd x index number and an even y index number.

Therefore, for index set K_2 , we have

$$\begin{aligned}
 L_{(1,2)} &= \{(0, 2), (4, 2), (8, 2)\}, \\
 R_{(1,2)} &= \{(2, 1), (2, 3), (6, 1), (6, 3)\}.
 \end{aligned}
 \tag{A.3}$$

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This study is supported by the Singapore Ministry of Education AcRF Tier 1 Grants RG117/14, M401030000 and the National Nature Science Foundation of China (Grant nos. 51338008, 11172035).

References

- [1] M. Abdulaal and L. J. LeBlanc, "Continuous equilibrium network design models," *Transportation Research Part B*, vol. 13, no. 1, pp. 19–32, 1979.
- [2] H. N. Tan, S. B. Gershwin, and M. Athans, *Hybrid Optimization in Urban Traffic Networks*, U.S. Transportation Systems Center, Cambridge, Mass, USA, 1979.
- [3] H.-J. Cho, *Sensitivity analysis of equilibrium network flows and its application to the development of solution methods for equilibrium network design problems [Ph.D. thesis]*, University of Pennsylvania, Philadelphia, Pa, USA, 1988.
- [4] T. L. Friesz, R. L. Tobin, H.-J. Cho, and N. J. Mehta, "Sensitivity analysis based heuristic algorithms for mathematical programs with variational inequality constraints," *Mathematical Programming*, vol. 48, no. 1–3, pp. 265–284, 1990.
- [5] H. Yang, "Sensitivity analysis for the elastic-demand network equilibrium problem with applications," *Transportation Research Part B: Methodological*, vol. 31, no. 1, pp. 55–70, 1997.
- [6] C. Suwansirikul, T. L. Friesz, and R. L. Tobin, "Equilibrium decomposed optimization: a heuristic for the continuous equilibrium network design problem," *Transportation Science*, vol. 21, no. 4, pp. 254–263, 1987.
- [7] H. Poorzahedy and M. A. Turnquist, "Approximate algorithms for the discrete network design problem," *Transportation Research Part B*, vol. 16, no. 1, pp. 45–55, 1982.
- [8] R. S. Solanki, J. K. Gorti, and F. Southworth, "Using decomposition in large-scale highway network design with a quasi-optimization heuristic," *Transportation Research B: Methodological*, vol. 33, no. 2, pp. 127–140, 1998.
- [9] Z. Gao, J. Wu, and H. Sun, "Solution algorithm for the bi-level discrete network design problem," *Transportation Research Part B: Methodological*, vol. 39, no. 6, pp. 479–495, 2005.
- [10] P. Luatthep, A. Sumalee, W. H. K. Lam, Z.-C. Li, and H. K. Lo, "Global optimization method for mixed transportation network design problem: a mixed-integer linear programming approach," *Transportation Research Part B: Methodological*, vol. 45, no. 5, pp. 808–827, 2011.
- [11] W. Y. Szeto and H. K. Lo, "Transportation network improvement and tolling strategies: the issue of intergeneration equity," *Transportation Research Part A*, vol. 40, no. 3, pp. 227–243, 2006.
- [12] H. K. Lo and W. Y. Szeto, "Time-dependent transport network design under cost-recovery," *Transportation Research B: Methodological*, vol. 43, no. 1, pp. 142–158, 2009.
- [13] H. Yang and M. G. H. Bell, "Models and algorithms for road network design: a review and some new developments," *Transport Reviews*, vol. 18, no. 3, pp. 257–278, 1998.
- [14] R. Z. Farahani, E. Miandoabchi, W. Y. Szeto, and H. Rashidi, "A review of urban transportation network design problems," *European Journal of Operational Research*, vol. 229, no. 2, pp. 281–302, 2013.
- [15] R. E. Allsop, "Some possibilities for using traffic control to influence trip distribution and route choice," in *Proceedings of the 6th International Symposium on Transportation and Traffic Theory, University of New South Wales, Sydney, Australia, 26–28 August, 1974*, Elsevier, 1974.
- [16] N. H. Gartner, S. B. Gershwin, J. D. C. Little, and P. Ross, "Pilot study of computer-based urban traffic management," *Transportation Research B*, vol. 14, no. 1–2, pp. 203–217, 1980.
- [17] T. Van Vuren and D. Van Vliet, *Route Choice and Signal Control: The Potential for Integrated Route Guidance*, Ashgate Publishing Company, Brookfield, Vt, USA, 1992.
- [18] G. E. Cantarella, G. Pavone, and A. Vitetta, "Heuristics for urban road network design: lane layout and signal settings," *European Journal of Operational Research*, vol. 175, no. 3, pp. 1682–1695, 2006.
- [19] P. Marcotte, "Network optimization with continuous control parameters," *Transportation Science*, vol. 17, no. 2, pp. 181–197, 1983.
- [20] Y. Sheffi and W. B. Powell, "Optimal signal settings over transportation networks," *Journal of Transportation Engineering*, vol. 109, no. 6, pp. 824–839, 1983.
- [21] B. G. Heydecker and T. K. Khoo, "The equilibrium network design problem," in *Proceedings of the Conference on Models and Methods for Decision Support (AIRO '90)*, Sorrento, Italy, 1990.
- [22] H. Yang and S. Yagar, "Traffic assignment and signal control in saturated road networks," *Transportation Research A*, vol. 29, no. 2, pp. 125–139, 1995.
- [23] Z.-Q. Luo, J.-S. Pang, D. Ralph, and S.-Q. Wu, "Exact penalization and stationarity conditions of mathematical programs with equilibrium constraints," *Mathematical Programming*, vol. 75, no. 1, pp. 19–76, 1996.
- [24] F. Teklu, A. Sumalee, and D. Watling, "A genetic algorithm approach for optimizing traffic control signals considering routing," *Computer-Aided Civil and Infrastructure Engineering*, vol. 22, no. 1, pp. 31–43, 2007.
- [25] E. Cascetta, M. Gallo, and M. Montella, "Optimal signal setting on traffic networks with stochastic equilibrium assignment," in *Proceedings of the Triennial Symposium on Transportation Analysis (TRISTAN III '98)*, San Juan, Puerto Rico, 1998.
- [26] C. Lee and R. B. Machemehl, "Genetic algorithm, local and iterative searches for combining traffic assignment and signal control," in *Proceedings of the Conference on Traffic and Transportation Studies*, pp. 488–497, Beijing, China, July 1998.

- [27] E. Cipriani and G. Fusco, "Combined signal setting design and traffic assignment problem," *European Journal of Operational Research*, vol. 155, no. 3, pp. 569–583, 2004.
- [28] S.-W. Chiou, "Bilevel programming for the continuous transport network design problem," *Transportation Research B: Methodological*, vol. 39, no. 4, pp. 361–383, 2005.
- [29] R. L. Tobin and T. L. Friesz, "Sensitivity analysis for equilibrium network flow," *Transportation Science*, vol. 22, no. 4, pp. 242–250, 1988.
- [30] A. Sumalee, D. P. Watling, and S. Nakayama, "Reliable network design problem: case with uncertain demand and total travel time reliability," *Transportation Research Record*, vol. 1964, pp. 81–90, 2006.
- [31] R. D. Connors, A. Sumalee, and D. P. Watling, "Sensitivity analysis of the variable demand probit stochastic user equilibrium with multiple user-classes," *Transportation Research B: Methodological*, vol. 41, no. 6, pp. 593–615, 2007.
- [32] J. Q. Ying, H. Lu, and J. Shi, "An algorithm for local continuous optimization of traffic signals," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1189–1197, 2007.
- [33] Q. Meng, H. Yang, and M. G. H. Bell, "An equivalent continuously differentiable model and a locally convergent algorithm for the continuous network design problem," *Transportation Research Part B: Methodological*, vol. 35, no. 1, pp. 83–105, 2001.
- [34] H. Ceylan and M. G. H. Bell, "Traffic signal timing optimisation based on genetic algorithm approach, including drivers' routing," *Transportation Research Part B: Methodological*, vol. 38, no. 4, pp. 329–342, 2004.
- [35] C. Lee and R. B. Machemehl, *Combined Traffic Signal Control and Traffic Assignment: Algorithms, Implementation and Numerical Results*, University of Texas at Austin, Austin, Tex, USA, 2005.
- [36] T. L. Friesz, H.-J. Cho, N. J. Mehta, R. L. Tobin, and G. Anandalingam, "Simulated annealing approach to the network design problem with variational inequality constraints," *Transportation Science*, vol. 26, no. 1, pp. 18–26, 1992.
- [37] T. L. Friesz, G. Anandalingam, N. J. Mehta, K. Nam, S. J. Shah, and R. L. Tobin, "The multiobjective equilibrium network design problem revisited: a simulated annealing approach," *European Journal of Operational Research*, vol. 65, no. 1, pp. 44–57, 1993.
- [38] H. Poorzahedy and O. M. Rouhani, "Hybrid meta-heuristic algorithms for solving network design problem," *European Journal of Operational Research*, vol. 182, no. 2, pp. 578–596, 2007.
- [39] F. Li, Z. Gao, K. Li, and D. Z. W. Wang, "Train routing model and algorithm combined with train scheduling," *Journal of Transportation Engineering*, vol. 139, no. 1, pp. 81–91, 2013.
- [40] D. Z. W. Wang and H. K. Lo, "Global optimum of the linearized network design problem with equilibrium flows," *Transportation Research B: Methodological*, vol. 44, no. 4, pp. 482–492, 2010.
- [41] H. Liu and D. Z. W. Wang, "Global optimization method for network design problem with stochastic user equilibrium," *Transportation Research Part B: Methodological*, vol. 72, pp. 20–39, 2015.
- [42] D. Z. W. Wang, H. Liu, and W. Szeto, "A novel discrete network design problem formulation and its global optimization solution algorithm," *Transportation Research Part E: Logistics and Transportation Review*, vol. 79, pp. 213–230, 2015.
- [43] R. Riemann, D. Z. W. Wang, and F. Busch, "Optimal location of wireless charging facilities for electric vehicles: flow-capturing location model with stochastic user equilibrium," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 1–12, 2015.
- [44] W. Y. Szeto, Y. Jiang, D. Z. W. Wang, and A. Sumalee, "A sustainable road network design problem with land use transportation interaction over time," *Networks and Spatial Economics*, pp. 1–32, 2013.
- [45] G. Ziyue and S. Yifan, "A reserve capacity model of optimal signal control with user-equilibrium route choice," *Transportation Research Part B: Methodological*, vol. 36, no. 4, pp. 313–323, 2002.
- [46] J. G. Wardrop, "Some theoretical aspects of road traffic research," *Proceedings of the Institution of Civil Engineers*, vol. 1, no. 3, pp. 325–362, 1952.
- [47] H. K. Lo, "A novel traffic signal control formulation," *Transportation Research Part A: Policy and Practice*, vol. 33, no. 6, pp. 433–448, 1999.
- [48] H. K. Lo, "A cell-based traffic control formulation: strategies and benefits of dynamic timing plans," *Transportation Science*, vol. 35, no. 2, pp. 148–164, 2001.
- [49] K. L. Croxton, B. Gendron, and T. L. Magnanti, "A comparison of mixed-integer programming models for nonconvex piecewise linear cost minimization problems," *Management Science*, vol. 49, no. 9, pp. 1268–1273, 2003.
- [50] A. B. Keha, I. R. de Farias Jr., and G. L. Nemhauser, "Models for representing piecewise linear cost functions," *Operations Research Letters*, vol. 32, no. 1, pp. 44–48, 2004.
- [51] J. Lee and D. Wilson, "Polyhedral methods for piecewise-linear functions I: the lambda method," *Discrete Applied Mathematics*, vol. 108, no. 3, pp. 269–285, 2001.
- [52] H. D. Sherali, "On mixed-integer zero-one representations for separable lower-semicontinuous piecewise-linear functions," *Operations Research Letters*, vol. 28, no. 4, pp. 155–160, 2001.
- [53] J. P. Vielma, S. Ahmed, and G. Nemhauser, "Mixed-integer models for nonseparable piecewise-linear optimization: unifying framework and extensions," *Operations Research*, vol. 58, no. 2, pp. 303–315, 2010.
- [54] F. Webster, *Traffic Signal Settings*, 1958.
- [55] M. J. Todd, "Union jack triangulations," in *Fixed Points: Algorithms and Applications*, S. Karamardian and C. B. Garcia, Eds., pp. 315–336, Academic Press, New York, NY, USA, 1977.
- [56] J. P. Vielma and G. L. Nemhauser, "Modeling disjunctive constraints with a logarithmic number of binary variables and constraints," *Mathematical Programming*, vol. 128, no. 1, pp. 49–72, 2011.
- [57] J. Breckman, "Encoding circuit," US Patent 2733432 A, 1956.
- [58] J. Löfberg, "YALMIP: a toolbox for modeling and optimization in MATLAB," in *Proceedings of the IEEE International Symposium on Computer Aided Control System Design*, pp. 284–289, IEEE, Taipei, Taiwan, September 2004.
- [59] IBM ILOG, *User's Manual for Cplex*, 2009, ftp://public.dhe.ibm.com/software/websphere/ilog/docs/optimization/cplex/ps_usrmanplex.pdf.
- [60] T. Larsson and M. Patriksson, "An augmented lagrangean dual algorithm for link capacity side constrained traffic assignment problems," *Transportation Research Part B*, vol. 29, no. 6, pp. 433–455, 1995.
- [61] S. Nguyen and C. Dupuis, "An efficient method for computing traffic equilibria in networks with asymmetric transportation costs," *Transportation Science*, vol. 18, no. 2, pp. 185–202, 1984.