A Macroscopic Signal Optimization Model for Arterials Under Heavy Mixed Traffic Flows

Yen-Yu Chen and Gang-Len Chang, *Member, IEEE*

***Abstract*—This paper presents a generalized signal optimization model for arterials experiencing multiclass traffic flows. Instead of using conversion factors for nonpassenger cars, the proposed model applies a macroscopic simulation concept to capture the complex interactions between different types of vehicles from link entry and propagation, to intersection queue formation and discharging. Since both vehicle size and link length are considered in modeling traffic evolution, the resulting signal timings can best prevent the queue spillback due to insufficient bay length and the presence of a high volume of transit or other types of large vehicles. The efficiency of the proposed model has been compared with the benchmark program TRANSYT-7F under both passenger flows only and multiclass traffic scenarios from modest to saturated traffic conditions. Using the measures of effectiveness of the average-delay-per-intersection approach and the total arterial throughput during the control period, our extensive numerical results have demonstrated the superior performance of the proposed model during congested and/or multiclass traffic conditions. The success of the proposed model offers a new signal design method for arterials in congested downtowns or megacities where transit vehicles constitute a major portion of traffic flows.**

***Index Terms*—Arterial control, multiclass traffic, signal optimization.**

I. Introduction

BEST using roadway capacity to contend with daily recurrent congestion has long been a priority task of urban traffic professionals. Over the past several decades, both researchers and practitioners have devoted tremendous resources to tackle this vital issue with various signal control strategies, such as arterial signal progression and responsive or proactive signal optimization systems. The emergence of intelligent transportation systems (ITS) since 1990s has also advanced the urban signal research to the real-time control level, based on multisource information available from the infrastructure and network vehicles. However, despite the immense potential of using advanced technologies to better capture traffic conditions, most cities in both developed and developing countries still depend on pretimed signal strategies to control their daily traffic congestion.

In review of the literature, it is noticeable that most existing studies on the subject of arterial signal optimization fall into one of the following two categories: the mathematical programming approach and the simulation-based method. Some pioneering models on the former category focused on maximizing arterial progression bandwidth with a set of mixed-integer linear programming formulations, giving less attention to flow interruptions caused by heavy or unbalanced turning volume [1]–[7]. More recent studies along the same line mainly addressed on producing signal timings to minimize the total arterial intersection delay [8], [9]. Despite the significant contribution of those models in the literature, the nature of mathematical programming formulations has limited their flexibility to capture complex intersection flow relations, such as the overflow from the tuning bay and the mutual lane blockage between through and turning lanes.

The latter category of studies aimed at capturing the complex interactions between traffic flows and their evolution under different signal phases using a simulation-based approach. TRANSYT-7F is one of the most well-recognized models in this category, which applies macroscopic traffic-flow formulations to model interrelations between arterial vehicle evolution and intersection queue formation. A near-optimal set of signal timing plans can be then generated via the imbedded search algorithms based on a predefined performance index [10], [11]. This simulation-based methodology has also been extended to perform online responsive signal optimization, based on real-time detected traffic data from available sensors. Examples of such systems include Sydney Coordinated Adaptive Traffic System (SCATS) [12], Split, Cycle, and Offset Optimization (SCOOT) [13], OPAC [14], and RHODES [15].

Using a similar concept but with different mathematical formulations, traffic control researchers have also produced various arterial signal optimization strategies, including store-and-forward models [16]–[18], queue-dispersion models [19]–[21], stochastic models [22], and discrete-time kinematic models [23]–[25]. Most of such studies, however, addressed mainly the undersaturated traffic conditions, giving less attention to the queue spillback and lane-blockage issues when the arterial links or turning bays become a potential congestion-contributing factor.

In parallel with the given studies for undersaturated arterials, some traffic researchers have developed various models for oversaturated intersections by distributing the delays caused

Manuscript received May 27, 2013; revised August 22, 2013 and October 11, 2013; accepted October 22, 2013. Date of publication November 26, 2013; date of current version March 28, 2014. The Associate Editor for this paper was J. E. Naranjo.

Y.-Y. Chen is with the Department of Transportation Technology and Management, National Chiao Tung University, Hsinchu 30050, Taiwan (e-mail: ymm887@yahoo.com.tw).

G.-L. Chang is with the Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: gang@umd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2013.2289961

by an oversaturated link to several intersections in the same arterial. For example, some researchers [26]–[28] developed a "bang–bang" control process to identify the optimal switching point for signal timings among intersections in an arterial. Their models were extended later to contend with various oversaturated intersections using flow propagation relationships [29], [30]. Along the same line, Abu-Lebdeh and Benekohal [31] modeled the intersection queue formation and dissipation at oversaturated links with a set of dynamic discrete formulations, and produced the optimal signal solutions with genetic algorithms (GA) [32] that can distribute excessive queues to multiple intersections over multiple cycles. Abu-Lebdeh *et al.* [33] later also presented a model that can capture the interactions of traffic throughput between neighboring intersections.

To tackle the complex oversaturated traffic issues, researchers for TRANSYT-7F [34] introduced a penalty function to the queue in its performance index if blockages between lanes occur. This enhanced version also enables users to have the option of using multiple cycles or a stepwise simulation program to account for spillback affects. Recently released versions of TRANSYT [35] has addressed this issue with a cell transmission model that offers more realistic illustration of the time-varying queue blocking effects between neighboring lanes and intersections. Using the same simulation-based concept but with different macroscopic formulations for traffic evolution and flow interactions, many researchers explored various models to capture dynamic queue interactions between neighboring intersections that subsequently affect the queue discharging rate and the resulting green time needs [36]–[39].

Despite the promising progress made by those studies in the literature on arterial signal optimization for both undersaturated and oversaturated traffic conditions, most existing models focused on signal control for arterials comprising mainly of passenger-car traffic and assumed that a typically small percentage of nonpassenger-car flows can be converted to equivalent passenger-car units (PCUs) [40], [41]. This type of conversion method even with volume-dependent factors (McTrans, 2008), however, may significantly underestimate the intersection queue and delay if traffic volume consists of a large percentage of nonpassenger-car flows, such as buses or commercial vehicles, as in many cities where transit vehicles constitute a major portion of urban commuting traffic. The effects of large vehicles on the delay, queue formation, link spillback, and lane blockage would be particularly pronounced when the link length between neighboring intersections is relative short and the turning bay is likely to be insufficient due to the right-of-way constraints, which are quite common in many urban downtowns in both developed and developing countries.

Grounded on the same methodology used by [39], this paper presents a generalized signal optimization approach for congested urban arterials that need to accommodate multiple classes of vehicles in traffic flows. The proposed model takes into account the operational characteristics and space needs of each vehicle class in formulating the traffic evolution, queue formation, and dissipation, allowing the optimized signal plan to more efficiently process multiclass traffic flows over urban congested arterials.

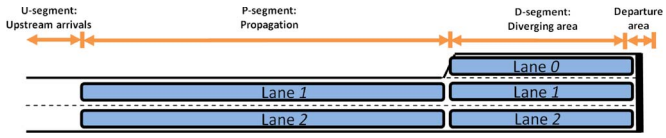

Fig. 1. Link decomposition.

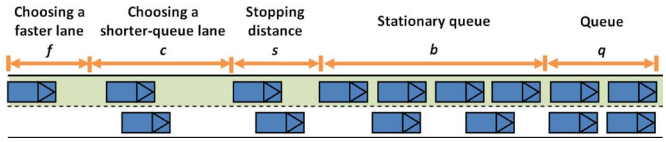

Fig. 2. P-segment and its subsegments.

II. Modeling Traffic Evaluation Dynamics in an Intersection Link

Consider a typical intersection approach, as shown in Fig. 1, which consists of one left-turn bay and two travel lanes. To capture the key traffic activities that may affect the discharging rate of an intersection and, consequently, the required signal times, one may divide each approach link into the following four segments for arriving, propagation (denoted as P-segment), diverging, and departure. Based on the potential lane-changing maneuvers within the propagation segment and the queue evolution in each departure lane, one can further define the P-segment into the following five subsegments (see Fig. 2):

- q subsegment: the length of the shorter queue between two through lanes;
- b subsegment: the difference between two queue lengths in those two through lanes;
- s subsegment: the stopping distance for vehicles evolving from the moving to the standing queue status;
- c subsegment: the travel distance for drivers from perceiving the intersection queue to join the moving queue;
- f subsegment: the travel distance within which drivers from the upstream link have not yet been affected by the intersection queue condition.

The first four segments are proposed to represent the sequential responses of drivers from entering f subsegment to join the intersection queues. For instance, drivers may take the distance of c subsegment to perceive the queue conditions in the available travel lanes and select a shorter queue. Depending on the approaching speed, vehicles may travel over the length of s subsegment to evolve from moving to the stationary queue state.

Note that the length of each subsegment is dynamic in nature and varies with the time-dependent entering volume and departure rates. Some of those subsegments may not exist over those control intervals without standing queues. Moreover, note that, for convenience of model description, the length of each segment is measured with a PCU.

To reflect the discrepancies between different types of vehicles (e.g., passenger cars, bus, and trucks) in the link evolution and queue discharging process, the formulations derived below take into account the actual space occupied by each type of vehicles and their start-up delays in computing the signal timings. The key feature of the proposed model is to produce an optimized signal plan for arterials with heavy multiclass traffic flows.

A. Compute the Arriving Flow Rate From Upstream U-Segment to P-Segment

Modeling of traffic evolving from entry to departure starts from the demand in the arrival segment (i.e., U-segment). Let D^E be the total demand of type-E vehicles during the study period; therefore, its flow rate during the kth interval $D^{E,k}$ can be expressed as follows:

$$D^{E,k} = \left(\frac{D^E}{3600} * \Delta t\right) \tag{1}$$

where Δt is the duration of one sliced time interval.

Note that, depending on the available physical space, some vehicles from (1) may not be able to move onto P-segment during interval k. Moreover, note that subscript "P" will be removed from the following formulations to compress the notation.

Let n_i^k and x_i^k denote the number of moving and queue vehicles (measured with the unit of average passenger car), respectively, in lane i of P-segment at step k; and C_i is the total storage space in link i of segment P. Then, its remaining space RS_i^k available to accommodate coming vehicles can be shown with

$$\mathrm{RS}_i^k = \left(C_i - n_i^k - x_i^k\right). \tag{2}$$

Let V_i^E represent the ratio of type-E vehicles over the total entering flow rate, and O_E denotes the ratio of the occupied space between a type-E vehicle and a passenger car. Then, with the assumption that every vehicle has the same opportunity to share the available space, the remaining space in lane i of p-segment during the kth interval to be allocated to type-E vehicles shall follow its proportion in the total flow rate (i.e., V_i^E), and the resulting number of such vehicles can be approximated as follows:

$$S_i^{E,k} = \frac{1}{o_E} * \left(C_i - n_i^k - x_i^k\right) * V_i^E. \tag{3}$$

Based on the demand from (1) and the available space shown in (3), the actual number of type-E vehicles $y_i^{E,k}$ to move onto lane i of P-segment during the kth interval shall be as follows:

$$y_i^{E,k} = \min\left\{D^{E,k}, S_i^{E,k}\right\}. \tag{4}$$

B. Modeling Lane-Changing Maneuvers and Traffic Evolution in Subsegment b

The second part of the model is to formulate the lane-changing maneuvers for turning and through vehicles in subsegments b, c, and f, which are based on the following steps.

Estimating Right-Turning Vehicles Within a Subsegment Which can Successfully Change Lanes: With the assumption that right-turning vehicles within subsegment b shall intend to change to the rightmost lane, each class of vehicles that can successfully make such changes depends certainly on its demand level and the space available on the target lane.

For the demand level, let $x_{1,b}^{E,k}$ be the total number of type-E right-turning vehicles queued at lane 1 of subsegment b in P-segment at step k, and $R_{1,b}^{E,R,k}$ denote the right-turn ratio out of the total type-E vehicles. Then, the number of these vehicles, i.e., $X_{1,b}^{E,R,k}$, which may have the desire to change lanes, can be expressed as follows:

$$X_{1,b}^{E,R,k} = x_{1,b}^{E,k} * R_{1,b}^{E,R,k}. \tag{5}$$

Certainly, depending on the available space in the rightmost lane, some vehicles from (5) may not be able to make necessary lane changes. For convenience of illustration, let lane 1 be assumed to have a longer queue, and the traffic condition in its neighbor lane (i.e., lane 2) is expected to be in a near-moving queue condition. Hence, the average space and time headways for vehicles in lane 2 within the b subsegment can be approximated as follows.

Average space headway in lane 2:

$$W_{2,b}^k = \frac{L_{2,b}^k - \Sigma_E\left(n_{1,b}^{E,k} * l^E\right)}{\Sigma_E n_{1,b}^{E,k}}. \tag{6}$$

Average time headway:

$$H_{2,b}^k = \frac{1}{v_{2,b}^k} * W_{2,b}^k \tag{7}$$

where $v_{2,b}^k$ and $n_{1,b}^{E,k}$ are the average speed of all vehicles and the number of type-E vehicles, respectively, in subsegment b of lane 2; $L_{2,b}^k$ denotes the total length of subsegment b; and l^E is the average length of a type-E vehicle.

Note that headways in subsegment b of lane 2 are assumed to distribute uniformly because those vehicles are likely in the forced-flow conditions and are ready to stop. Moreover, note that the resulting average headway (or space headway) may not serve the need of all vehicle types due to the discrepancy in their vehicle sizes. Hence, the following binary indicator is adopted to reflect such a discrepancy:

$$I_2^{E,b,k} = \begin{cases} 1 & \text{if } H_{2,b}^k \geq t^E \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where t^E denotes the minimum headway needed for type-E vehicles to change lanes.

Based on the assumption that the percentage of each type of vehicles in the lane-changing flows from lane 1 is proportional to its distribution in the total flow, one can approximate the ratio of type-E vehicles in the total lane-changing flows from lane 1 as follows:

$$M_{1,b}^{E,R,k} = \frac{R_{1,b}^{E,R,k} * I_2^{E,b,k}}{\Sigma_E\left(R_{1,b}^{E,R,k} * I_2^{E,b,k}\right)} \tag{9}$$

where $R_{1,b}^{E,R,k}$ denotes the ratio of type-E right-turning vehicles over its total flows at subsegment b in lane 1 at step k.

Hence, by assuming that all vehicles intending to change lanes regardless of their differences in size have the same probability to occupy the available space in the target lane, then the number of type-E vehicles $S_{2,b}^{E,R,k}$ allowed to share the total remaining space in subsegment b in the target lane at

step k for lane changes can be represented with the following expression:

$$S_{2,b}^{E,R,k} = \frac{1}{O_E} * M_{1,b}^{E,R,k} * \left(C_{2,b}^{k} - n_{2,b}^{k}\right) \tag{10}$$

where $n_{2,b}^{k}$ and $C_{2,b}^{k}$ are the number of existing vehicles and the total available space, respectively, at subsegment b in lane 2 of segment P at step k, and O_E is the average length of one type-E vehicle.

Based on the actual lane-changing demand from (5) and the available space for such changes from (10), the number of type-E right-turning vehicles $y_{1,2}^{E,b,R,k}$ at subsegment b in lane 1, which can successfully move to lane 2, is naturally the minimum of (10) and (5) as shown in the following:

$$y_{1,2}^{E,b,R,k} = \min\left\{X_{1,b}^{E,R,k}, S_{2,b}^{E,R,k}\right\}. \tag{11}$$

If $X_{1,b}^{E,R,k} > S_{2,b}^{E,R,k}$, $S_{2,b}^{E,R,k}$ vehicles can successfully change lanes, and $(X_{1,b}^{E,R,k} - S_{2,b}^{E,R,k})$ vehicles will not be able to make lane changes at step k. Those vehicles may change lanes at the next time step.

Estimating the Through Vehicles Changing From Lane 1 to Lane 2: By assuming that the queue length in lane 1 is b feet longer than its neighboring lane (i.e., lane 2), it is expected that some through vehicles may change to lane 2. Hence, one can follow the same procedures [i.e., (5) –(11)] to approximate the number of through vehicles by type, i.e., $S_{2,b}^{E,T,k}$, that can perform the lane changes based on the allocated space as follows:

$$S_{2,b}^{E,T,k} = \frac{1}{O_E} * \frac{R_{1,b}^{E,T,k} * I_2^{E,b,k}}{\Sigma_E\left(R_{1,b}^{E,T,k} * I_2^{E,b,k}\right)} * \left(C_{2,b}^{k} - n_{2,b}^{k} - y_{1,2}^{E,b,R,k}\right). \tag{12}$$

Note that the given equation assumes that right-turning vehicles due to the mandatory nature have the high priority to use the space in the rightmost lane. Similarly, based on the lane-changing demand and available space, the number of type-E through vehicles $y_{1,2}^{E,b,T,k}$ that can successfully change to the lane of a shorter queue at interval k can be estimated with the following:

$$y_{1,2}^{E,b,T,k} = \min\left\{X_{1,b}^{E,T,k}, S_{2,b}^{E,T,k}\right\}. \tag{13}$$

C. Modeling Lane-Changing Maneuvers and Traffic Evolution in Subsegments c and f

Estimate Right-Turning Vehicles Making Lane Changes in Subsegment c: As shown in Fig. 2, vehicles entering subsegment c are generally able to perceive the queue lengths on the current and neighboring lanes. Assuming that drivers are likely to change to the lane of shorter queue unless constrained by the turning need, such a demand $N_{1,c}^{E,R,k}$ by vehicle type naturally varies with the following variables:

- the number of type-E moving vehicles at subsegment c in lane 1 of segment P at step k, i.e., $n_{1,c}^{E,k}$;
- the ratio of right-turning type-E vehicles over the total flows at subsegment c in lane 1 of segment P at step k, i.e., $R_{1,c}^{E,R,k}$;
- the probability that the headways at time k in the receiving lane are larger than the critical gap for type-E vehicles to change lanes, i.e., $P_{2,c}^{E,k}$.

The following illustrates their relations with the number of type-E right-turning moving vehicles $N_{1,c}^{E,R,k}$ at subsegment c in lane 1 of segment P at step k, which may intend to perform lane-changing activities:

$$N_{1,c}^{E,R,k} = P_{2,c}^{E,k} * n_{1,c}^{E,k} * R_{1,c}^{E,R,k} \tag{14}$$

where $P_{2,c}^{E,k} = P(T \geq t^E)$.

Note that vehicles can successfully change lanes only if there are acceptable gaps in the target lane. A precise modeling of such a lane-changing process is a complex stochastic queuing issue of moving service windows versus time-varying demand and may be excessively tedious for signal control need if the concerns are mainly on the pretimed cycle length and green times. Hence, one can select a proper and simple distribution based on field data to approximate the probability used in the given equation.

By the same procedure for modeling traffic evolution on subsegment b, one shall first compute the remaining space to be allocated to different types of right-turning vehicles, which is $(C_{2,c}^{k} - n_{2,c}^{k})$, where $C_{2,c}^{k}$ denotes the total space in lane 2 of subsegment c at time interval k, and $n_{2,c}^{k}$ shows the occupied link space during the same interval.

Assuming that each vehicle has the same probability of sharing the available space, then the total space to be occupied by type-E right-turning vehicles in lane 1, i.e., $N_{1,c}^{E,R,k}$, should be equal approximately to its volume ratio over the total right-turning flow rate. The following shows such an approximation:

$$S_{2,c}^{E,R,k} = \frac{1}{O_E} * \frac{N_{1,c}^{E,R,k}}{\Sigma_E N_{1,c}^{E,R,k}} * \left(C_{2,c}^{k} - n_{2,c}^{k}\right) \tag{15}$$

where $S_{2,c}^{E,R,k}$ is the number of type-E right-turning vehicles that are allowed to change to lane 2 based on the allocated link space. Hence, the actual number of type-E vehicles $y_{1,2}^{E,c,R,k}$ on subsegment c that can successfully change to lane 2 shall be the minimum of (14) and (15), and be expressed as follows:

$$y_{1,2}^{E,c,R,k} = \min\left\{N_{1,c}^{E,R,k}, S_{2,c}^{E,R,k}\right\}. \tag{16}$$

Estimating Lane-Changing Through Vehicles on Subsegment c: Unlike right-turning vehicles, drivers in lane 1 may consider changing to lane 2 only if they may encounter a shorter queue. Thus, following the same procedures for turning vehicles, one can approximate the number of through vehicles $N_{1,c}^{E,T,k}$, which may consider lane changes as follows:

$$N_{1,c}^{E,T,k} = P_{2,c}^{E,k} * n_{1,c}^{E,k} * R_{1,c}^{E,T,k} * I_{1,2}^{c,k} \tag{17}$$

where

$$I_{1,2}^{c,k} = \begin{cases} 1 & \text{if lane 2 has a shorter queue} \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

Similarly, by assuming that right-turning vehicles will be given the priority to change to the rightmost lane (i.e., lane 2), the remaining space for through vehicles $S_{2,c}^{E,T,k}$ in subsegment c that are allowed to change lanes can be approximated as follows:

$$S_{2,c}^{E,T,k} = \frac{1}{O_E} * \frac{N_{1,c}^{E,T,k}}{\Sigma_E N_{1,c}^{E,T,k}} * \left(C_{2,c}^{k} - n_{2,c}^{k} - y_{1,2}^{E,c,R,k}\right). \quad (19)$$

By the same token, the number of through vehicles in subsegment c that can successfully change lanes shall be the minimum of

$$y_{1,2}^{E,c,T,k} = \min\left\{N_{1,c}^{E,T,k}, S_{2,c}^{E,T,k}\right\}. \quad (20)$$

Modeling Lane-Changing Maneuvers and Traffic Evolution in Subsegment f: Since drivers in subsegment f are assumed to behave similarly in subsegments c, except that their lane-changing decisions would be based on increasing the speed rather than reducing the encountered queue. Hence, one can apply the given derivation procedures to approximate the number of vehicles by type in each lane in subsegment f that may change lanes at each time interval. The binary indicator with (18) should, however, be changed as follows:

$$I_{1,2}^{f,k} = \begin{cases} 1 & \text{if lane 2 has a higher speed} \\ 0 & \text{otherwise.} \end{cases} \quad (21)$$

D. Modeling of the Traffic Evolution From the Propagation to Diverge Segments

As shown in Fig. 1, the number of vehicles by type that can flow out from segment P to the diverging segment D conceivably depends on the following factors:

- the number of vehicles in segment P;
- the maximum outgoing flow rate from each lane in segment P;
- the available space in each receiving lanes in segment D to accommodate different types of vehicles.

Note that a potential lane blockage is likely to occur and needs to be taken into account in computing the outgoing flow rate if the receiving lane in segment D is a turning bay.

Number of Vehicles Available to Move From Segment P to Segment D: Notably, the traffic evolution in lane 1 from the propagation to diverging segments is more complex than in lane 2 because its neighboring turning bay may experience overflow if having an excessive number of left-turning vehicles. Hence, the following derivation starts with the traffic relations in lane 2.

Since segment P comprises subsegments f, b, and s, it takes the following time for a vehicle in lane 2 to evolve over the entire segment and move onto the diverging segment D:

$$T_2^{E,k} = \frac{L_{2,f}^{k}}{v_{2,f}^{E,k}} + \frac{L_{2,c}^{k}}{v_{2,c}^{E,k}} + \frac{L_{2,b}^{k}}{v_{2,b}^{E,k}} + \frac{L_{2,s}^{k}}{v_{2,s}^{E,k}} \quad (22)$$

where

$T_2^{E,k}$ total travel time through subsegments f, b, and s of lane 2 at time k;

$L_{2,j}^{k}$ length of subsegment j in lane 2 of segment P at step k ($j = f, c, b, s$);

$v_{2,j}^{E,k}$ The average speed of subsegment j in lane 2 of segment P at step k ($j = f, c, b, s$).

The last two terms in the given equation will be zero if no queue exists in both Lanes 1 and 2. For notation compression, let the time-varying travel time in (22) be represented as follows:

$$r(k) = \left\lfloor \frac{T_2^{E,k}}{\Delta t} \right\rfloor. \quad (23)$$

Hence, those vehicles, i.e., $y_2^{E,k}$ entering lane 2 in segment P between time intervals $(k - r(k))$ and k, will be too far from moving into segment D prior to the kth interval. A mathematical expression of those vehicles is shown in the following:

$$y_2^{E,k} = \sum_{t=k-r(k)}^{k} y_2^{E,t}. \quad (24)$$

The following shows the number of type-E vehicles in lane 2 eligible to flow into segment D, i.e., $N_2^{E,k}$, which equals the summation of those in the queue $x_2^{E,k}$ and moving states $n_2^{E,k}$, and less those $y_2^{E,k}$ by (24), as follows:

$$N_2^{E,k} = \left(n_2^{E,k} - y_2^{E,k} + x_2^{E,k}\right). \quad (25)$$

Given these vehicles ready to move into the diverging segment, one can now proceed to analyze the actual number of vehicles that can be discharged to the receiving lane in segment D.

Note that, to approximate the outgoing flow rates under the given mixed-flow density at time k, different types of vehicles are assumed to travel at different speeds under the same traffic condition. For example, passenger cars and motorcycles can generally move faster than buses in the same arterial link. Thus, given the $V_2^{E,k}$ ratio of type-E vehicles out of the total mixed-flow density of ρ_2^k, one can apply a precalibrated speed–density relationship to compute the speed of each vehicle type under the mixed-flow density as $v_2^{E,k}$, and approximate the total potential outgoing flow rate for type-E vehicle as $(V_2^{E,k} * \rho_2^k) * v_2^{E,k}$.

However, since given the existence of a left-turn bay in the discharging segment D, only the through and right-turning vehicles will move to its through lane (i.e., lane 2); the number of type-E vehicles available to move onto lane 2 can be stated as follows:

$$F_2^{E,k} = \frac{1}{O_E} * \left(R_2^{E,T,k} + R_2^{E,R,k}\right) * \left(V_2^{E,k} * \rho_2^k * v_2^{E,k}\right) * \Delta t \quad (26)$$

where $R_2^{E,T,k}$ and $R_2^{E,R,k}$ are the through and right-turn ratio of type-E vehicles over the total type-E vehicles in lane 2 of segment P at step k; $v_2^{E,k}$ denotes the average speed of type-E vehicles over the same segment and time interval; and $V_2^{E,k}$ is the ratio of type-E vehicles over the total flow in lane 2.

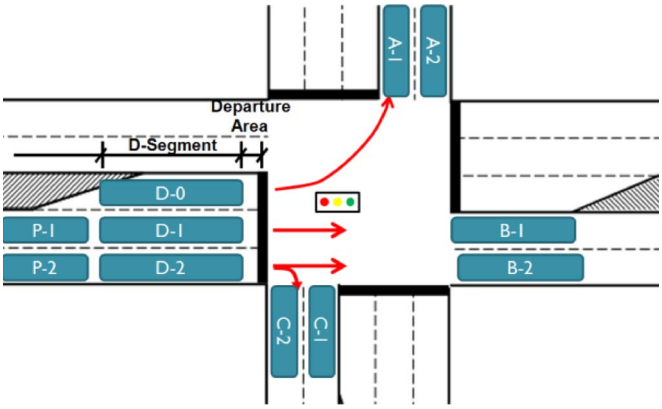

Fig. 3. Discharging process from segment D.

Conceivably, by the same assumption that each type of vehicles has the same probability to share the available space in the receiving lane of segment D, the space in the receiving lane for type-E vehicles should equal their ratio in the total outgoing flow rate, as shown in

$$J_2^{E,k} = \frac{v_2^{E,k} * \rho_2^k * v_2^{E,k} * \Delta t}{\Sigma_E v_2^{E,k} * \rho_2^k * v_2^{E,k} * \delta t}. \tag{27}$$

One can then approximate the number of type-E vehicles in lane 2 allowed to move from segment P into segment D, based on allocated space, at interval k as follows:

$$S_2^{E,k} = \frac{J_2^{E,k}}{O_E} * \left(C_{D,2} - n_{D,2}^k - x_{D,2}^k\right) \tag{28}$$

where

$C_{D,2}$ storage space of lane 2 in segment D;
$n_{D,2}^k$ number of moving vehicles in lane 2 of segment D at step k;
$x_{D,2}^k$ number of queuing vehicles in lane 2 of segment D at step k.

Based on the supply (25), demand (26), and traffic conditions (28), one can compute the actual number of vehicles in lane 2 that can successfully move into the diverging segment D as follows:

$$y_{2,D,2}^{E,k} = \min\left\{N_2^{E,k},\ F_2^{E,k},\ S_{D,2}^k\right\}. \tag{29}$$

Note that the given procedures for lane 2 vehicles are applicable to approximate the evolution of lane 1 traffic flow, except that their receiving lanes in segment D could be either lane 1 or lane 0, i.e., the left-turn bay (see Fig. 3).

If lane 1 in segment D has been fully occupied by different types of vehicles, the turning bay would not be able to accommodate any left-turning vehicles. The following is proposed to define the status of such a blockage:

$$Q_{D,1}^k = \begin{cases} 0 & \text{if } \left(C_{D,1} - n_{D,1}^k - x_{D,1}^k\right) = 0 \\ 1 & \text{if } \left(C_{D,1} - n_{D,1}^k - x_{D,1}^k\right) > 0 \end{cases} \tag{30}$$

where $Q_{D,1}^k$ is set to zero if lane 1 of segment D has been fully occupied by through vehicles at interval k. Then, the available space for each vehicle type $S_1^{E,k}$ from lane 1 to the receiving lanes in segment D can be expressed as the following general form:

$$S_1^{E,k} = \frac{J_1^{E,k}}{O_E} * \left(C_{D,1} - n_{D,1}^k - x_{D,1}^k\right) * Q_{D,1}^k. \tag{31}$$

Note that the given formulations assume that the left-turn bay length is sufficiently long to accommodate left-turning vehicles per cycle. Otherwise, the overflows from the left-turn bay may block the through lanes and cause complex mutual blockage issues. Nevertheless, one can still apply the same derivation process [39] to contend with those intersection-queue-blockage-related issues.

Computing the Discharging Flow Rate From Segment D: As shown in Fig. 3, the number of vehicles to be physically discharged from each lane in segment D at time interval k to lane i in the receiving link R depends on the following three variables:

- number of vehicles by type present in each lane in segment D;
- maximum outgoing flow rate of type-E vehicles in each lane;
- remaining space to be occupied by type-E vehicles in the receiving lane.

With respect to the total number of vehicles in segment D eligible to discharge to any downstream receiving lane, one can estimate it with the following

$$N_{D,z}^{E,k} = \left(n_{D,z}^{E,k} - y_{D,z}^{E,k} + x_{D,z}^{E,k}\right) * U_{R,i}^E, \quad z = 0, 1, 2 (\text{in segment } D) \tag{32}$$

where $n_{D,z}^{E,k}$ and $x_{D,z}^{E,k}$ are the number of type-E moving and queue vehicles in lane z of segment D at step k, respectively; $y_{D,z}^{E,k}$ denotes the number of type-E vehicles entering the same segment between time interval $(k - r(k))$ and k; and $U_{R,i}^E$ denotes the lane usage ratio of lane i in the receiving link R that reflects drivers' preference of taking each available receiving lane.

For the maximum number of vehicles by type that can move out of segment D at time k, it, as discussed previously, would be the product of the following two variables:

- the total number of vehicles that can be discharged from segment D during the sliced interval and the signal status, i.e., $Q_{D,z}^E * \Delta t * G_p^{H,k}$, where $Q_{D,z}^E$ is the discharging rate, and $G_p^{H,k}$ is the green phase indicator (0 or 1) of phase p at intersection H at step k;
- the ratio of each type of vehicles over the total flow rate in segment D's discharging flow rate, i.e., $(V_{D,z}^{E,k} * \rho_{D,z}^k * v_{D,z}^{E,k} * \Delta t)/(\sum_E v_{D,z}^{E,k} * \rho_{D,z}^E * v_{D,z}^{E,k} * \Delta t)$, where $\rho_{D,z}^k$ is the density of lane z $(z = 0, 1, 2)$ in segment D at step k, and $v_{D,z}^{E,k}$ is the average speed of type-E vehicles in lane z of segment D at step k.

Again, for notation compression, the subscript denoting segment D is removed from the questions here.

The following shows the product of these two variables:

$$F_z^{E,k} = \frac{v_z^{E,k} * \rho_z^k * v_z^{E,k} * \Delta t}{\Sigma_E v_z^{E,k} * \rho_z^E * v_z^{E,k} * \Delta t} * Q_z^E * \Delta t * G_p^{H,k} * U_{R,i}^E. \quad (33)$$

Note that (33) assumes that each type of vehicles has its own speed–density relation with respect to the mixed-flow density, and the discharging rate will follow the ratio of each type of vehicles in the total flow rate.

With the same notion as used earlier, one can also approximate the remaining space in the receiving lane to be occupied by type-E vehicles as follows:

$$S_{A,i}^{E,k} = \frac{1}{O_E} * \frac{v_z^{E,k} * \rho_z^k * v_z^{E,k} * \Delta t}{\Sigma_E \left\{ V_z^{E,k} * \rho_z^E * v_z^{E,k} * \Delta t \right\}} * \left(C_{A,i} - n_{A,i}^k - x_{A,i}^k \right) \quad (34)$$

where $n_{A,i}^k$ and $x_{A,i}^k$ denote the number of moving and queue vehicles, respectively, to the receiving lanes. Note that by taking into account the available space in the receiving lane, the proposed model can account for the impact of potential downstream oversaturated condition in the design of signal control where vehicles cannot be discharged during the green phase and the intersection may experience queue spillback.

In brief, one can follow the same procedure to estimate the number of vehicles that can move from each lane in segment D to their target receiving lane at time interval k as follows:

$$y_{z,A,i}^{E,k} = \min \left\{ N_z^{E,k},\, F_z^{E,k},\, S_{A,i}^{E,k} \right\}. \quad (35)$$

To complete the set of constraints for arterial signal control, (36)–(40) are specified to reflect the typical signal relations, and (41) is used to reflect the flow conservation relations in each intersection link, as shown in the following:

$$g_{h1} + g_{h2} + g_{h3} + g_{h4} = C_h \quad (36)$$

$$g_{hj} \geq G_{\min,j} = 1, \ldots, 4 \quad (37)$$

$$C_{\min} \leq C_h \leq C_{\max} \quad (38)$$

$$\Sigma g_{h_j} + \Sigma I_{h_j} = C_h \quad (39)$$

$$GI_p^{H,k} = \begin{cases} 1 & \text{if } \sum_{j=1}^{p-1} (G_{Hj} + I_{Hj}) < \operatorname{mod}(k * \Delta t, C) \\ & \quad \leq \sum_{j=1}^{p-1} (G_{Hj} + I_{Hj}) + G_{Hp} \\ 0 & \text{otherwise} \end{cases} \quad (40)$$

$$n_{i,j}^{E,k+1} + x_{i,j}^{E,k+1} = n_{i,j}^{E,k} + x_{i,j}^{E,k} + \sum_{l,m} y_{l,m,i,j}^{E,k} - \sum_{o,n} y_{i,j,o,n}^{E,k} \quad (41)$$

where

- $n_{i,j}^{E,k+1}$ number of type-E moving vehicles in lane j of segment i at step $k+1$;
- $x_{i,j}^{E,k+1}$ number of type-E queuing vehicles in lane j of segment i at step $k+1$;
- $n_{i,j}^{E,k}$ number of type-E moving vehicles in lane j of segment i at step k;
- $x_{i,j}^{E,k}$ number of type-E queuing vehicles in lane j of segment i at step k;
- $y_{i,m,i,j}^{E,k}$ number of type-E vehicles that is from lane m of segment l to lane j of segment i at step k;
- $y_{i,j,o,n}^{E,k}$ number of type-E vehicles that is from lane j of segment i to lane n of segment o at step k.

III. Formulations of the Control Objective Functions

Consider the presence of $(n+1)$ vehicles in the queue at time interval k and the total discharge time for the $(n+1)^{th}$ vehicle shall be equal to the sum of the following two components:

- sum of the expected discharging time for the first n vehicles; and
- expected discharging time of the $(n+1)^{th}$ vehicle.

Assuming that traffic flows comprise only two types of vehicles and that there are $B_{n,i}^k$ buses and $P_{n,i}^k$ passenger cars in the first n vehicles of queue, one can then approximate their total delay as follows:

$$B_{n,i}^k * T_B + P_{n,i}^k * T_P \quad (42)$$

where T_B and T_P are the average discharging time per bus and per passenger car, respectively.

Due to the discharge time discrepancy between a bus and a passenger car from a standing queue position, the critical issue in (42) is to compute the probability of having B_n^k buses and P_n^k passenger cars in the first n queue vehicles from the given total flow rate.

Let $\binom{x_{D,i}^k}{n}$ be the number of possible sequences for n queue vehicles out of the $x_{D,i}^k$ queue vehicles in queue at lane i of segment D at step k. Then, the probability of having exactly $B_{n,i}^k$ buses and $P_{n,i}^k$ passenger cars in its first n vehicles of queue shall be as follows:

$$\frac{\binom{x_{D,i}^{B,k}}{B_{n,i}^k} * \binom{x_{D,i}^{P,k}}{P_{n,i}^k}}{\binom{x_{D,i}^k}{n}} \quad (43)$$

where

- $B_{n,i}^k$ total number of buses in the first n vehicles at lane i of segment D at step k;
- $x_{D,i}^{B,k}$ total number of buses in queue at lane i of segment D at step k;
- $P_{n,i}^k$ total number of buses in the first n vehicles at lane i of segment D at step k;
- $x_{D,i}^{P,k}$ total number of passenger cars in queue at lane i of segment D at step k;
- n first n vehicles in queue in lane i of segment D at step k;
- $x_{D,i}^k$ total number of vehicles in queue at lane i of segment D at step k.

Hence, if only $x_{D,i}^{P,k}$ passenger cars and $x_{D,i}^{B,k}$ buses are in the traffic queue, one can sum up the probability of all possible

combinations B_n^k to compute the total expected discharging time of the first n vehicles in the queue as follows:

$$\sum_{B_n^k} \frac{\binom{x_{D,i}^{B,k}}{B_{n,i}^k} * \binom{x_{D,i}^{P,k}}{P_{n,i}^k}}{\binom{x_{D,i}^k}{n}} * \left[B_{n,i}^k * T_B + P_{n,i}^k * T_p\right]. \quad (44)$$

Since the probability for the $(n+1)^{th}$ vehicle to be a bus or a passenger car can be approximated as $(x_{D,i}^{B,k} - B_{n,i}^k)/(x_{D,i}^k - n)$ and $(x_{D,i}^{P,k} - P_{n,i}^k)/(x_{D,i}^k - n)$, respectively, the additional discharging time needed for the $(n+1)^{th}$ vehicle can be shown as follows:

$$\frac{\left(x_{D,i}^{B,k} - B_{n,i}^k\right)}{x_{D,i}^k - n} * T_B + \frac{\left(x_{D,i}^{P,k} - P_{n,i}^k\right)}{x_{D,i}^k - n} * T_p. \quad (45)$$

Thus, the summation of (44) and (45) offers a reasonable approximation of the required total expected time $W_{D,i,n+1}^k$ to discharge $(n+1)$ queue vehicles at lane i of segment D at step k.

Since the total delay equals the sum of expected discharging times of all vehicles plus their waiting time at the red phase, one can show the objective function of minimizing the total delay as follows:

$$\min \quad \sum_i \sum_k \left(\left(\sum_n W_{D,i,n}^k \right) + \Delta t * I_{D,i}^k * x_i^k \right) \quad (46)$$

where $I_{D,i}^k$ is the red time indicator for lane i in segment D at step k, and x_i^k is the number of queuing vehicles at interval k in lane i.

The alternative objective function of maximizing the total throughput is relatively straightforward and can be show with

$$\max \quad \sum_{E,k,i} y_{D,i,\text{out}}^{E,k} \quad (47)$$

where $y_{D,i,\text{out}}^{E,k}$ denotes discharged type-E vehicles at lane i in subsegment D at step k.

IV. Solution Algorithm

As in most optimization models for network signals, the formulations for multiple vehicle classes, including both binary variables and nonlinear system constraints, are difficult to find in the global optimal solution. To be efficient for use in practice, this paper has applied two heuristic methods and selected the one that yields better signal settings as the solution.

The first one proposed by Yue and Chang [39] is a GA-based method that can produce a reasonably effective solution after an extensive experimental process. As shown in Fig. 4, to generate control parameters that satisfy the signal optimization constraints, the proposed heuristic applies the following decoding scheme based on the phase structure. First, let NP_n be the number of phases at intersection n. A total number of $\text{NP}_n + 1$ fractions will be then generated for the controller at intersection n from the decomposed binary strings by converting the binary string to a decimal number and dividing the number by the maximum possible decimal number represented by the binary string. The $\text{NP}_n + 1$ fractions are used to code the green times and cycle length, as shown in the following equations:

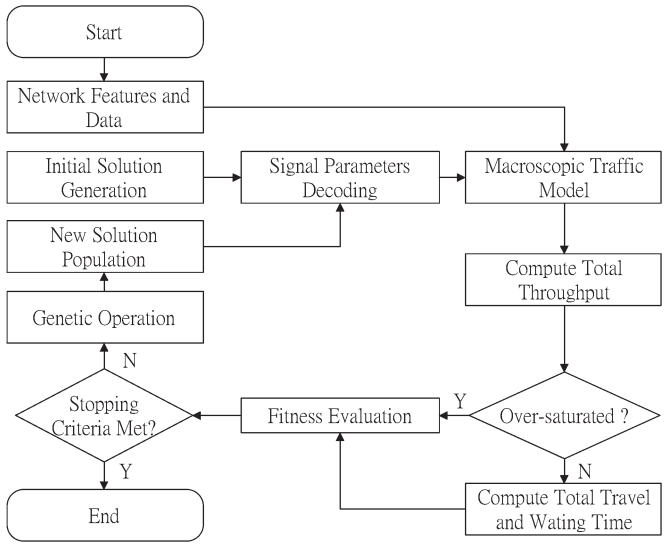

Fig. 4. Flowchart of the solution algorithm with the proposed GA method [39].

$$G_{np} = G_{np}^{\min} + \left(C - \sum_{j \in P_n} G_{nj}^{\min} - \sum_{j \in P_n} I_{nj} \right) * \lambda_p * \prod_{j=1}^{p} (1 - \lambda_{j-1}),$$
$$p = 1, \ldots, \text{NP}_n - 1; \quad n \in S_N \quad (48)$$

$$G_{np} = G_{np}^{\min} + \left(C - \sum_{j \in P_n} G_{nj}^{\min} - \sum_{j \in P_n} I_{nj} \right) * \prod_{j=1}^{p} (1 - \lambda_{j-1}),$$
$$p = \text{NP}_n; \quad n \in S_N \quad (49)$$

where

$C_{\min}, C_{\max}$	minimal and maximal cycle lengths;
$G_{np}^{\min}$	minimal green time for phase p of intersection n;
C	common cycle length for the target arterial for the given period T;
G_{np}	green time for phase p at intersection n for the given period T;
I_{np}	intergreen time for phase p at intersection n.

The first population of GA will be generated randomly, and each individual will be decoded to a set of signal timing plans based on the aforementioned scheme. Then, one can use the proposed network flow model to compute the objective value for the given analysis period T. The corresponding fitness measure can be obtained from the objective function value. Based on the fitness evaluation, the crossover and mutation procedures will be executed and continued until reaching the

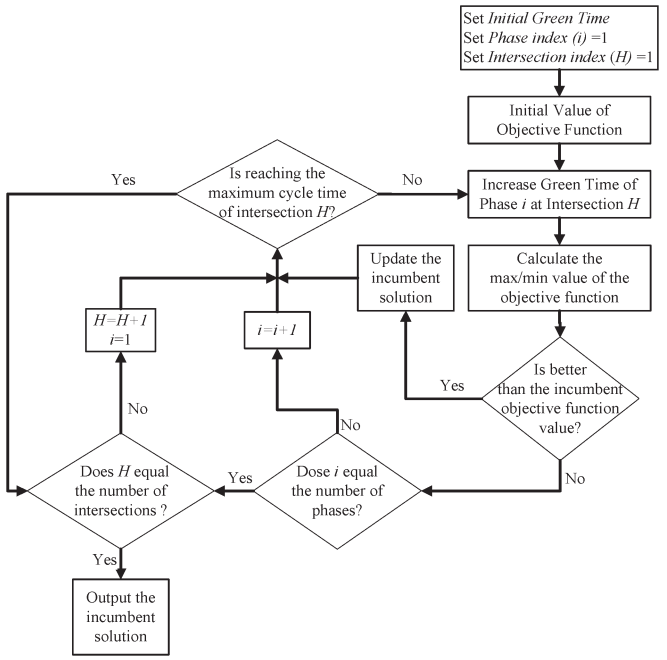

Fig. 5. Flowchart of the solution algorithm with the proposed gauzy branch-and-bound method.

stop criterion. Note that such a solution algorithm will first maximize the total system throughput and then switch to minimize the total travel time in the arterial network if the traffic conditions are undersaturated (see Fig. 4). The characteristics of GA are available from the literature [39].

The second algorithm, as shown in Fig. 5, is a gauzy branch-and-bound method that uses an incremental search method to sequentially identify a better signal setting and cycle length, which is based on a preselected set of initial values and maximal number of iterations. With this heuristic, the initial green times are first set to the minimum green times, and phase index i and intersection index H are set to 1. Then, the initial value of the objective function can be computed.

The search procedure is to extend the green time of phase i at intersection H at an increment of one second each time, followed by recomputing the new objection function. If the new value for the objective function is better than the incumbent one, then the solution will be updated. After updating with the solution, the algorithm will examine whether the current cycle time has reached its upper bound (the maximum cycle time). The green time of phase i at intersection H will be increased again if the current cycle length is below the upper bound. If the upper bound has been reached and the intersection index does not equal the number of intersections, then it will increase one unit to the intersection index.

If the new value of the objective function is not better than the incumbent objective function value, the algorithm will increase the green time for another phase at the current intersection. If all phases of the current intersection have been calculated, the green time of the first phase at the next intersection shall be added one increment per second. If one of the following rules has been matched, the algorithm shall stop the search process.

1) i and H reach the upper bounds of the phase index and the intersection index, respectively.

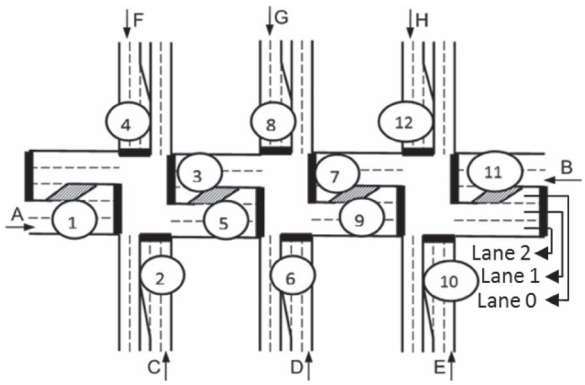

Fig. 6. Network of three intersections and 12 approaches for numerical analysis.

TABLE I
KEY MODEL PARAMETERS AND GEOMETRIC CHARACTERISTICS USED IN THE EXPERIMENTAL SCENARIOS

Parameters	
Free flow speed	64.4 kph(40 mph)
Jam density	130.4 veh/km/lane (210 veh/mile/lane)
Passenger car length	4.57 m (15 ft)
Bus length	9.14 m (30 ft)
Max cycle length	150 s
Minimal green time	10 s (phase 1, phase 3) 15 s (phase 2, phase 4)
Inter-green time	3 s
Saturation flow rate	1600 vph/lane
Link length	372 m (1220 feet)
Number of lanes	2 through lane / 1 left turn bay
Left turn bay length	64 m (210 feet)

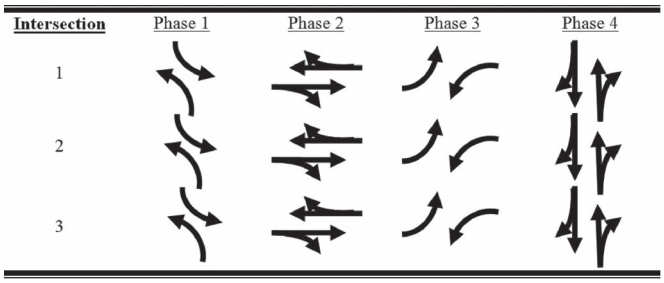

Fig. 7. Phase settings.

2) Cycle times of all intersections have reached the maximum.

Due to the complex nature of the signal optimization model for arterials experiencing multiclass traffic flows, neither solution method can guarantee to reach the global optimal value after reaching the preset maximum iteration. Although the proposed solution algorithm cannot guarantee to obtain the global optima, our experimental process has identified that the following search sequence can yield a better solution: 1) starting the search from the highest demand intersection in a descending order to the one with the lowest demand in a network; and 2) at each intersection, searching from the approach having the highest inflows in a descending order to the one with the lowest traffic volume. Hence, the converged solutions from both approaches are used to compare with the benchmark program TRANSYT-7F to evaluate their efficiency

TABLE II
EXPERIMENTAL SCENARIOS FOR MODEL EVALUATION

Scenario	Objective function	Mixed Flow	Bus Ratio	Demand entries (vph/lane)							
				A	B	C	D	E	F	G	H
1	Min	No	0	600	600	500	500	500	500	500	500
2	Min	No	0	700	700	500	500	500	500	500	500
3	Min	No	0	800	800	500	500	500	500	500	500
4	Max	No	0	600	600	500	500	500	500	500	500
5	Max	No	0	700	700	500	500	500	500	500	500
6	Max	No	0	800	800	500	500	500	500	500	500
7	Max	No	0	900	900	500	500	500	500	500	500
8	Max	No	0	1000	1000	500	500	500	500	500	500
9	Min	Yes	0.3	600	600	500	500	500	500	500	500
10	Min	Yes	0.3	700	700	500	500	500	500	500	500
11	Max	Yes	0.3	600	600	500	500	500	500	500	500
12	Max	Yes	0.3	700	700	500	500	500	500	500	500
13	Max	Yes	0.3	800	800	500	500	500	500	500	500
14	Max	Yes	0.3	900	900	500	500	500	500	500	500
15	Max	Yes	0.3	1000	1000	500	500	500	500	500	500

and effectiveness. Since our experimental results indicate that the gauzy branch-and-bound method outperform the GA methods in most scenarios, the following discussion is focused on comparing the performance between the gauzy branch-and-bound method and TRANSYT-7F.

V. NUMERICAL ANALYSIS

To ensure the effectiveness of the proposed mixed-flow model, this paper has conducted a performance comparison with TRANSYT-7F under various undersaturated and oversaturated traffic conditions with and without multiple classes of vehicles. Note that the performance indicators for comparison include throughput and the average delay in each individual intersection and at the network level. All such results from each scenario under both control models are produced from the simulation results with CORSIM using the average of 15 replications. Each simulation replication runs 1.5 h, and the simulation results of first 30 min have been removed from the output analysis due to the system stability concern. Fig. 6 illustrates the network of three intersections and 12 approaches for numerical analysis, where some key experimental parameters and geometric characteristics listed in Table I were based on the field survey results from the city of Hsinchu, Taiwan. The phasing plans for these intersections are shown in Fig. 7.

Table II presents 15 scenarios for evaluation, including the flow rate by the approach, the control objective function, and the bus ratio over the total traffic flow.

Note that scenarios 1–3, containing passenger cars only, illustrate relatively moderate traffic conditions, and their optimal signal solutions were generated with the objective function of minimizing the total vehicle delay. In contrast, maximizing the total throughput was used as the control objective for scenarios 4 to 8 that show the congested and saturated traffic conditions. Hence, comparing the performance with TRANSYT-7F under these passenger-car only scenarios can verify the effectiveness of the core traffic evolution logic embedded in our proposed mixed-flow model.

Scenarios 9–15 present the same traffic scenarios as in scenarios 1–8 but with 33% of buses in the total traffic volume. The performance comparison under these scenarios is selected to further evaluate the effectiveness of the proposed model under mixed-flow scenarios.

Fig. 8 shows the comparison of delays in each approach between TRANSYT-7F and our proposed model under all scenarios. Table III presents the average delay difference in percentage per approach under all scenarios; Table IV summarizes the discrepancy in the total network delay in each scenario; and Table V shows the total throughput for those scenarios under the maximum throughput objective function.

For relatively moderate traffic conditions under scenarios 1 and 2, the proposed model clearly outperforms the benchmark program with respect to both the average delay in each intersection approach (see Fig. 8) and the average network delay per vehicle (see Table IV). When the demand in both primary entry nodes (i.e., nodes A and B) are increased to a higher level, as shown in scenario 3, the results reveal that the average vehicle delays in each approach and at the network level are both less than that with TRANSYT-7F, except in these two primary entry nodes. This reflects the key feature of the proposed model, which, by accounting for link and vehicle lengths, can produce signal timing plans to minimize the likelihood of having overflows in the primary inner network links.

Scenarios 4–8 are designed to investigate the performance of the proposed model under congested and single-mode traffic conditions. Using the maximizing the throughput as the control objective, Table V presents the resulting throughput during the control period, showing the trend of throughput difference (compared with the benchmark) increasing with the demand volume. Further comparison based on the delays at the network level and in each approach also confirms that the proposed model has better performance. For example, under scenario 8, the produced model yields the average network delay of 3.26 min/vehicle, about 36% less than with the compared model (see Table IV). It also produces 5.45% more of the total throughput (see Table V) and the lower average delay in 10 out of 12 intersection approaches (see Table IV and Fig. 8).

Scenarios 9–15 are used to evaluate the performance of the proposed model under heavy mixed flows (i.e., 30% of buses in the total demand) and from moderate to congested traffic conditions. First of all, with the proposed mixed-flow signal optimization model for scenarios 9 and 10, i.e., moderate

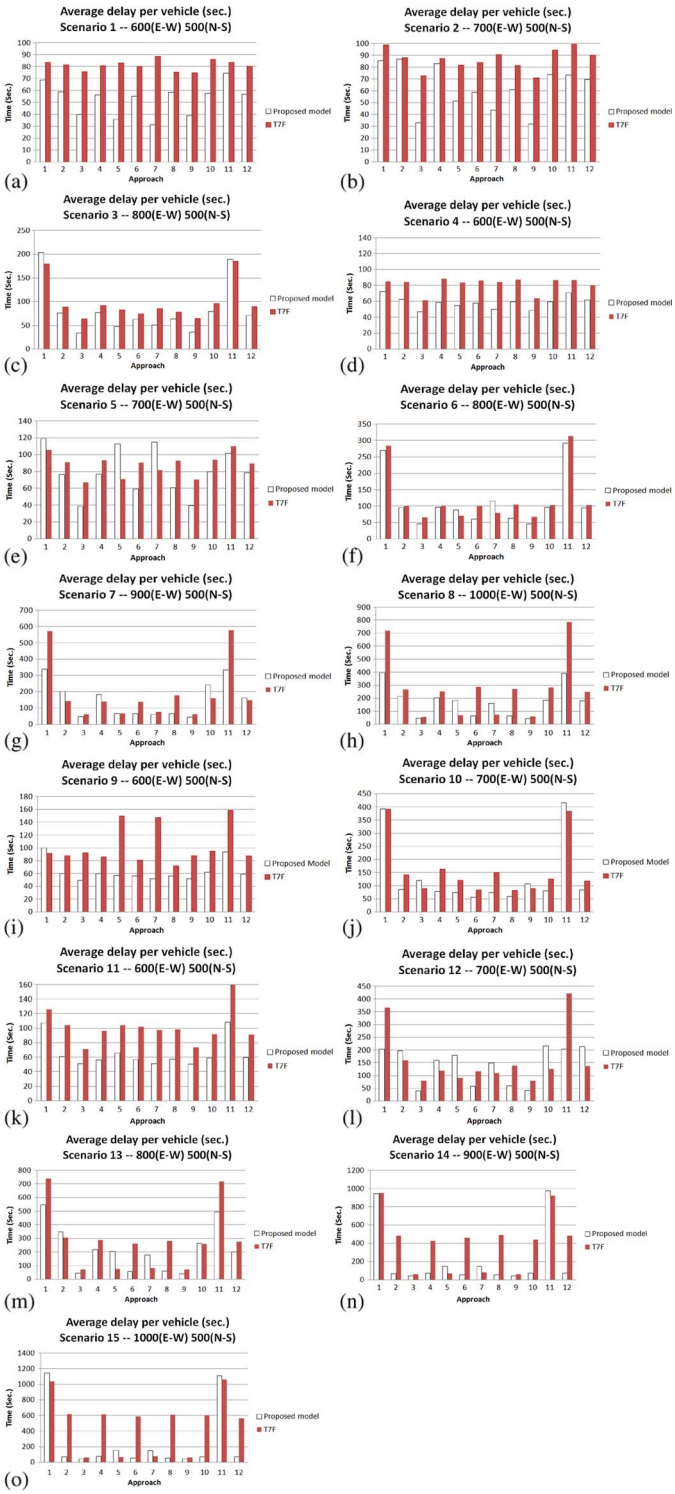

Fig. 8. Comparison of average delay per vehicle by the intersection approach.

mixed traffic flow conditions, the resulting average network delays of 1.06 and 2.41 min/vehicle are less than 1.73 and 2.83 min/vehicles, respectively, by the benchmark model (see Table IV). A further comparison of average delay in each intersection approach also confirms that the superior performance of the proposed model, where 11 and 10 out of 12 intersection approaches in scenarios 9 and 10, respectively, experience less delays. In scenarios 12–15, the delays of the proposed model in approaches 5 and 7 are higher than with TRANSYT-7F but not their downstream approaches (i.e., approaches 3 and 9, respectively). This is due likely to the discrepancy in their cycle lengths because both methods are designed to minimize the total network delay but not to optimize the offsets for maximal progression. Hence, the cycle length at each intersection varies with its demand level and the applied optimization method. The distribution of different cycle lengths among different intersections will result in different delay distributions among all intersection approaches even under the same total demand pattern.

By applying the same performance indicators to those congested mixed-flow conditions, the results in Table V indicate that both models yield the same level of throughput for scenarios 11–13 because all volumes are able to go through the network during the control time period. The resulting delays experienced by those vehicles under these two models, however, are quite different, where the proposed mixed-flow model yields much less average network vehicle delays, ranging from 14% to 36%, in all three scenarios than with the benchmark model, and produces more intersection approaches with less delays per vehicle (i.e., 12, 6, and 8 out of 12 approaches).

The benefit of the proposed model that considers vehicle size in the mixed traffic flow seems much more pronounced in the oversaturated mixed-flow conditions, as shown in scenarios 14 and 15. Notably, its average network delays per vehicle for scenarios 14 and 15 are 40.42% and 43.42%, respectively, which are less than that with TRANSYT-7F, but it produces 8.42 and 10.29%, respectively, which are more throughputs than under the same scenarios. With respect to the delay in each approach, the proposed model with its features of accounting for physical space need of different vehicle types has resulted in lower average delays in 8 and 9 approaches, respectively, out of 12 approaches in scenarios 14 and 15.

VI. Conclusion

This paper has presented a signal optimization model for urban arterials with heavy mixed traffic flows, such as buses or other types of transit vehicles, which is a quite common traffic flow pattern in major cities in developing countries. The proposed model, using a macroscopic traffic simulation concept, can account for the effects of each vehicle type's physical length on the traffic evolution and queue formation process, which offers an effective way to process mixed vehicle flows to go through series of intersections with the minimal likelihood of incurring lane blockage or link spillback. Although the proposed methodology remains exploratory in nature, the results of extensive numerical comparisons with a state-of-the-art signal optimization program has shown its promising properties in terms of minimizing the total delay and increasing the throughput under both passenger-car only and mixed traffic flow conditions.

Future extensions along this research line will be focused on the following areas: 1) incorporating some small vehicles such as motorcycles in the mixed traffic flow; 2) enhancing the solution heuristics to ensure the computing efficiency; and 3) overcoming the limitations of the current models. The primary limitations of the proposed model lie in its use of the

TABLE III
AVERAGE DELAY DIFFERENCE BETWEEN THE PROPOSED AND BENCHMARK MODELS

Unit: %

Approach	1	2	3	4	5	6	7	8	9	10	11	12
Scenario1	-17.3	-27.5	-47.0	-30.5	-56.9	-31.1	-64.8	-22.8	-47.8	-33.3	-11.0	-29.1
Scenario2	-13.7	-1.2	-54.6	-5.0	-37.1	-30.1	-51.7	-25.3	-55.1	-21.8	-26.0	-22.4
Scenario3	13.3	-14.0	-47.5	-16.4	-42.8	-15.7	-40.9	-18.7	-45.0	-17.9	2.2	-20.7
Scenario4	-15.3	-25.9	-23.6	-33.2	-34.3	-33.0	-40.7	-32.3	-24.2	-31.5	-18.8	-23.3
Scenario5	12.2	-15.9	-41.9	-17.3	59.5	-34.7	41.1	-34.5	-43.4	-14.8	-8.0	-12.7
Scenario6	-5.2	-3.6	-29.9	-3.5	25.4	-39.5	46.1	-40.5	-31.1	-6.2	-6.9	-8.5
Scenario7	-41.0	41.1	-25.6	30.1	-1.2	-52.7	-18.6	-62.8	-27.9	52.1	-41.9	11.5
Scenario8	-45.3	-18.8	-17.4	-18.1	168.3	-77.9	122.4	-75.7	-23.4	-34.1	-50.2	-27.0
Scenario9	9.4	-32.0	-46.8	-30.9	-62.2	-30.7	-65.0	-21.9	-40.9	-34.4	-40.8	-32.5
Scenario10	0.1	-39.3	34.1	-52.5	-38.7	-32.9	-50.2	-29.5	19.0	-35.7	8.5	-28.9
Scenario11	-14.6	-41.5	-28.7	-41.3	-36.7	-44.1	-47.2	-41.7	-31.1	-35.9	-39.9	-34.2
Scenario12	-44.6	24.4	-48.2	34.4	98.5	-50.8	37.1	-57.0	-47.4	72.1	-52.0	56.9
Scenario13	-26.1	13.2	-38.4	-23.7	171.7	-78.4	114.1	-79.1	-43.0	1.0	-31.0	-27.3
Scenario14	-1.1	-85.7	-26.1	-82.8	108.9	-87.7	81.0	-88.3	-29.6	-83.4	5.5	-84.9
Scenario15	10.3	-88.5	-18.8	-87.7	139.8	-90.4	99.5	-90.6	-26.1	-88.0	4.6	-87.5

TABLE IV
DIFFERENCE IN THE NETWORK VEHICLE DELAY BETWEEN THE PROPOSED AND BENCHMARK MODELS

Unit: minute

Scenario	The proposed model	T7F	Diff.
1	0.88	1.35	-34.88%
2	1.05	1.45	-28.03%
3	1.48	1.73	-14.39%
4	0.98	1.36	-27.95%
5	1.36	1.48	-7.68%
6	2.06	2.23	-7.98%
7	2.65	3.60	-26.19%
8	3.26	5.15	-36.76%
9	1.06	1.73	-38.74%
10	2.41	2.83	-15.05%
11	1.11	1.74	-36.24%
12	2.42	2.83	-14.35%
13	3.91	5.04	-22.35%
14	4.30	7.21	-40.42%
15	4.87	8.61	-43.42%

TABLE V
HOURLY THROUGHPUT BETWEEN THE PROPOSED AND BENCHMARK MODELS

Unit: Number of Vehicles

Scenario	The proposed model	T7F	Diff.*
4	8421	8395	0.31%
5	8768	8770	-0.02%
6	9008	9004	0.04%
7	9121	9016	1.16%
8	9359	8875	5.45%
11	8440	8471	-0.37%
12	8499	8574	-0.87%
13	8490	8415	0.89%
14	8587	7920	8.42%
15	8642	7836	10.29%

* [(Total Throughput of the proposed model) - (Total Throughput of T7F)]/(Total Throughput of T7F)

average volume from field data as the input, which in reality, could be an interval and vary from day to day. Hence, robust optimization may be more effective than the currently used heuristics. Moreover, the assumption that drivers will take the mandatory lane change when available, instead of waiting to the last moment, may not always be consistent with the aggressive driving patterns. Moreover, the implicit assumption that all drivers with different vehicle types will behave identically under the same traffic queue and lane-changing environment shall be further released in the future work.

REFERENCES

[1] N. H. Gartner, J. D. C. Little, and H. Gabbay, "Optimization of traffic signal settings by mixed integer linear programming. Part I: The network coordination problem," *Transp. Sci.*, vol. 9, no. 4, pp. 321–343, Nov. 1975.

[2] N. H. Gartner, J. D. C. Little, and H. Gabbay, "Optimization of traffic signal settings by mixed-integer linear programming Part II: The network synchronization problem," *Transp. Sci.*, vol. 9, no. 4, pp. 344–363, Nov. 1975.

[3] N. H. Gartner, S. F. Assmann, F. Lasaga, and D. L. Hou, "MULTIBAND-A variable bandwidth arterial progression scheme," *Transp. Res. Rec.*, no. 1287, pp. 212–222, 1990.

[4] N. H. Gartner, S. F. Assmann, F. L. Lasaga, and D. L. Hou, "A multi-band approach to arterial traffic signal optimization," *Transp. Res. B*, vol. 25, no. 1, pp. 55–74, Feb. 1991.

[5] S. L. Cohen and C. C. Liu, "The bandwidth-constrained TRANSYT signal optimization program," *Transp. Res. Rec.*, no. 1057, pp. 1–9, 1986.

[6] N. A. Chaudhary and C. J. Messer, "PASSER-IV: a program for optimizing signal timing in grid networks," *Transp. Res. Rec.*, no. 1421, pp. 82–93, 1993.

[7] C. Stamatiadis and N. H. Gartner, "MULTIBAND-96: A program for variable bandwidth progression optimization of multiarterial traffic networks," *Transp. Res. Rec.*, no. 1554, pp. 9–17, 1996.

[8] S. C. Wong, "A lane-based optimization method for minimizing delay at isolated signal-controlled junctions," *J. Math. Model. Algorithms*, vol. 2, no. 4, pp. 379–406, Dec. 2003.

[9] Q. K. He, K. L. Head, and J. Ding, "PAMSCOD: Platoon-based arterial multi-modal signal control with online data," *Transp. Res. C*, vol. 20, no. 1, pp. 164–184, Feb. 2012.

[10] D. I. Robertson, "TRANSYT: A traffic network study tool," Road Res. Lab., Berkshire, U.K., RRL Rep. LR 253, 1969.

[11] C. E. Wallace, K. G. Courage, D. P. Reaves, G. W. Shoene, G. W. Euler, and A. Wilbur, *TRANSYT-7F User's Manual*. Gainesville, FL, USA: Univ. Florida, 1988.

[12] P. Lowrie, "The Sydney coordinated adaptive control system—Principles, methodology, algorithms," *IEEE Conf. Publication*, no. 207, pp. 67–70, 1982.

[13] D. I. Robertson and R. D. Bretherton, "Optimizing networks of traffic signals in real-time: the SCOOT method," *IEEE Trans. Veh. Technol.*, vol. 40, no. 1, pp. 11–15, Feb. 1991.

[14] N. H. Gartner, "OPAC: A demand-responsive strategy for traffic signal control," *Transp. Res. Rec.*, no. 906, pp. 75–81, 1983.

[15] P. Mirchandani and L. Head, "A real-time traffic signal control system: Architecture, algorithms, and analysis," *Transp. Res. C*, vol. 9, no. 6, pp. 415–432, Dec. 2001.

[16] M. G. Singh and H. Tamura, "Modeling and hierarchical optimization of oversaturated urban road traffic networks," *Int. J. Control*, vol. 20, no. 6, pp. 913–934, Dec. 1974.

[17] G. C. D'Ans and D. C. Gazis, "Optimal control of oversaturated store and forward transportation networks," *Transp. Sci.*, vol. 10, no. 1, pp. 1–19, Feb. 1976.
[18] M. Papageorgiou, "An integrated control approach for traffic corridors," *Transp. Res. C*, vol. 3, no. 1, pp. 19–30, Feb. 1995.
[19] H. R. Kashani and G. N. Saridis, "Intelligent control for urban traffic systems," *Automatica*, vol. 19, no. 2, pp. 191–197, Mar. 1983.
[20] J. Wu and G. L. Chang, "An integrated optimal control and algorithm for commuting corridors," *Int. Trans. Oper. Res.*, vol. 6, no. 1, pp. 39–55, Jan. 1999.
[21] A. Van den Berg, B. Hegyi, De Schutter, and J. Hellendoorn, "A macroscopic traffic flow model for integrated control of freeway and urban traffic networks," in *Proc. 42nd IEEE Int. Conf. Decision Control*, Maui, HI, USA, 2003, pp. 2774–2779.
[22] X.-H. Yu and W. Recker, "Stochastic adaptive control model for traffic signal systems," *Transp. Res. C*, vol. 14, pp. 263–282, Aug. 2006.
[23] H. Lo, "A novel traffic signal control formulation," *Transp. Res. A*, vol. 33, no. 6, pp. 433–448, Aug. 1999.
[24] H. Lo, "A cell-based traffic control formulation: Strategies and benefits of dynamic timing plans," *Transp. Sci.*, vol. 35, no. 2, pp. 148–164, May 2001.
[25] C. Beard and A. Ziliaskopoulos, "System optimal signal optimization formulation," *Transp. Res. Rec.*, no. 1978, pp. 102–112, 2006.
[26] P. G. Michalopoulos and G. Stephanopolos, "Oversaturated signal system with queue length constraints—I: Single intersection," *Transp. Res.*, vol. 11, no. 6, pp. 413–421, Dec. 1977.
[27] P. G. Michalopoulos and G. Stephanopolos, "Oversaturated signal systems with queue length constraints—II: Systems of intersections," *Transp. Res.*, vol. 11, no. 6, pp. 423–428, Dec. 1977.
[28] P. G. Michalopoulos and G. Stephanopolos, "Optimal control of oversaturated intersections: theoretical and practical considerations," *Traffic Eng. Control*, vol. 19, no. 5, pp. 216–221, Jan. 1978.
[29] T. Chang and J.-T. Lin, "Optimal signal timing for an oversaturated intersection," *Transp. Res. B*, vol. 34, no. 6, pp. 471–491, Aug. 2000.
[30] T.-H. Chang and G.-Y. Sun, "Modeling and optimization of an oversaturated signalized network," *Transp. Res. B*, vol. 38, no. 8, pp. 687–707, Sep. 2004.
[31] G. Abu-Lebdeh and R. F. Benekohal, "Development of a traffic and queue management procedure for oversaturated arterials," *Transp. Res. Rec.*, no. 1603, pp. 119–127, Dec. 1997.
[32] M. Girianna and R. F. Benekohal, "Using genetic algorithms to design signal coordination for oversaturated networks," *J. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 117–129, Apr. 2004.
[33] G. Abu-Lebdeh, H. Chen, and R. F. Benekohal, "Modeling traffic output for design of dynamic multi-cycle control in congested conditions," *J. Intell. Transp. Syst.*, vol. 11, no. 1, pp. 25–40, Jan. 2007.
[34] M.-T. Li and A. C. Gan, "Signal timing optimization for oversaturated networks using TRANSYT-7F," *Transportation Research Record*, no. 1683, pp. 118–126, 1999.
[35] J. C. Binning, G. Burtenshaw, and M. Crabtree, *TRANSYT 13 User Guide*. Berkshire, U.K.: Transport Research Laboratory, 2008.
[36] B. Park, C. J. Messer, and T. Urbanik, "Traffic signal optimization program for oversaturated conditions: Genetic algorithm approach," *Transp. Res. Rec.*, no. 1683, pp. 133–142, 1999.
[37] I. Yun and B. Park, "Application of stochastic optimization method for an urban corridor," in *Proc. Winter Simul. Conf.*, 2006, pp. 1493–1499.
[38] A. Stevanovic, P. T. Martin, and J. Stevanovic, "VisSim-based genetic algorithm optimization of signal timings," *Transp. Res. Rec.*, no. 2035, pp. 59–68, 2007.
[39] Y. Yue and G.-L. Chang, "An arterial signal optimization model for intersections experiencing queue spillback and lane blockage," *Transp. Res. C*, vol. 19, no. 1, pp. 130–144, Feb. 2011.
[40] D. Mariagrazia, P. Maria, and C. Meloni, "A signal timing plan formulation for urban traffic control," *Control Eng. Pract.*, vol. 14, no. 11, pp. 1297–1311, Nov. 2006.
[41] Y.-C. Chang and Y.-F. Huang, "Stepwise genetic fuzzy logic signal control under mixed traffic conditions," *J. Adv. Transp.*, vol. 47, no. 1, pp. 43–60, Jan. 2013.

Yen-Yu Chen received the B.S. degree in traffic and transportation management from Feng Chia University, Taichung, Taiwan, in 2003 and the M.S. degree in transportation technology and management from National Chiao Tung University, Hsinchu, Taiwan, in 2005. He is currently working toward the Ph.D. degree with the Department of Transportation Technology and Management, National Chiao Tung University.

His current research interests include network traffic control and advanced traveler information systems.

Gang-Len Chang (M'13) received the B.Eng. degree from National Cheng Kung University, Tainan, Taiwan, in 1975; the M.S. degree from National Chiao Tung University, Hsinchu, Taiwan, in 1979; and the Ph.D. degree in transportation engineering from the University of Texas at Austin, TX, USA, in 1985.

He is currently a Professor with the Department of Civil and Environmental Engineering, University of Maryland, College Park, MD, USA. His current research interests include network traffic control, freeway traffic management and operations, real-time traffic simulation, and dynamic control of urban systems.

Dr. Chang is a member of the American Society of Engineers (ASCE). He has served as the Chief Editor for the *ASCE Journal of Urban Planning and Development* over the past 13 years.