Modelling and Traffic Signal Control of Heterogeneous Traffic Systems

Yicheng Zhang, Student Member, IEEE, Rong Su, Senior Member, IEEE, Yi Zhang, Student Member, IEEE, and Chunyang Sun

Abstract—An urban traffic system is a heterogeneous system, which consists of different types of intersections and dynamics. In this paper, we focus on one type of heterogeneous traffic network, which consists of signalized junctions and non-signalized ones, where in the latter case vehicles usually follow the first-in-first-out principle. We propose a novel model describing the dynamic behaviors of such a system and validate it via simulations in VISSIM. Upon such a new model, a signal control problem for a heterogeneous traffic network is formulated as a mixed integer programming problem, which is solved by a Lagrangian multiplier based hierarchical distributed approach. Comparisons between a homogeneous traffic system and a heterogeneous one are provided, which leaves the door open for developing a systematic planning approach on deciding what traffic junctions require signal control to ensure a good traffic control performance, thus, have a great social and economic potentials, considering that it is rather expensive to have signal control in an urban area.

I. INTRODUCTION

Traffic congestions in urban area occur frequently, which affect daily life and pose all kinds of problems and challenges. Alleviation of traffic congestions not only improves travel safety and efficiencies but also reduces environmental pollution. The urban traffic system consists of intersections and links, which include controllers, flow dynamics and volume dynamics. The traffic system modelling and the traffic signal design are the two important phases in solving the urban traffic congestion problem.

From the system modelling side, the urban traffic system is a heterogeneous system consisting of multiple kinds of intersections, dynamics and participants. In general, the intersections in urban traffic system could be classified into two main groups, i.e., the signalized intersections and the non-signalized intersections. A signalized intersection equips a set of traffic signals while a non-signalized or uncontrolled intersection is one in which the entrance into the intersection from any of the approaches is not controlled by a regulatory sign or a traffic signal. Although it seems that putting traffic signals at all intersections in the urban area would lead to a very orderly fashion, the costs of building and maintaining signalized intersections are much higher than the costs of non-signalized intersections. Moreover, it may be unnecessary to signalize all the intersections because of the low traffic flow for some minor intersections. Thus in real-world practice, both signalized and non-signalized intersections are adopted in the urban traffic system, which lead to the heterogeneous features of the link volume dynamics and the traffic flow dynamics. A detailed survey on monitoring an intersection based on the behaviors of the participants is shown in [1]. It provides a microscopic view or Lagrangian Model of the traffic dynamics at the intersection. For a macroscopic view or Eulerian Model of the traffic system, some descriptions about the non-signalized intersections or uncontrolled intersections can be found in the literature [2]-[9].

From the traffic signal scheduling strategy design view, there are four different traffic signal control strategies, namely, fixed time strategies versus traffic responsive strategies, and isolated strategies versus coordinated strategies. Notable strategies proposed in the last few decades include, SIGSET [10], SIGCAP [11], TRANSYT [2] [12], SCOOT [13] [14], OPAC [15] and PRODYN [16]. From the system modelling view, Lighthill and Whitham in 1955 and Richards in 1956 provided a good foundation in macroscopic traffic system modelling, the LWR model [17] [18]. A breakthrough came when Daganzo developed a finite difference solution scheme for the LWR model by adopting a simplified Fundamental Diagram [19] [20], which he called the cell-transmission model (CTM). In [21] and [22], the author transforms CTM into a mixed integer linear programming (MILP) problem. However, the transformation is not equivalent rigorously. Moreover, the formulation does not include any analysis about the relationship among the speed, previous signal status and velocity. Another weak point is that in this model the cycle time is fixed and the signal scheduling strategy is based on the fixed-time cycle. Lin and Wang [23] define difference cell models in the traffic network and different constraint sets are proposed for each type of cells. In literature [24], an urban road traffic light control problem is formulated as a scheduling problem, aiming to reduce the total waiting time over a given finite horizon. One of the key contributions in the model is to describe each outgoing flow rate as a nonlinear mixed logical switching function over the source link volume, the destination link volume and capacity, and the drivers potential psychological response to the past traffic light signals. The outgoing flow rate model makes the proposed approach applicable to both under-saturated and over-saturated situations. The traditional concepts of cycles, splits, and offsets are not adopted in this framework, making the proposed approach fall in the class of model-based optimization methods, where each traffic light is
assigned with a green light period in a real-time manner by the network controller. Previously the traffic signal control (TSC) problem is solved by using fixed time strategy with individual or coordinated intersections to adjust the green signal ratio in the traffic system. It can be formulated as an optimization problem and solved by standard optimization methodologies or computational intelligence techniques [25].

In this paper, a novel model for the urban traffic system is proposed to describe the heterogeneous traffic system with both signalized intersections and non-signalized intersections. For the signalized intersections, the proposed model describes the flow dynamics for controllable traffic flows and uncontrollable traffic flows, e.g., some merging flows from the free-turning directions. For the non-signalized intersections, the model provides a method to estimate the flow dynamics based on the gap-acceptance model and first-come-first-serve principle. The proposed non-signalized intersection model is validated based on the simulations in VISSIM simulator, a microscopic traffic simulation platform [26]. Moreover, a signal control problem formulation for the heterogeneous traffic network is proposed as a mixed integer programming problem.

In the simulation section, some comparisons among the fully controlled traffic system (all intersections are signalized), partially controlled traffic system and fully uncontrolled traffic system (all intersections are non-signalized) are provided, which leaves the door open for developing a systematic planning approach on deciding what traffic junctions require signal control to ensure a good traffic control performance.

The contributions of this work are listed below. Firstly, we develop and validate a heterogeneous traffic network model with both signalized and non-signalized intersections; secondly, a traffic signal control problem is formulated based on this model to minimize the network-wise traffic delay; thirdly, we do some simulations to to provide some guidance for the design of urban traffic system.

The paper is organized as follows. In Section II, we introduce the urban traffic system model for the problem development. Mathematical models for both signalized intersections and non-signalized intersections are proposed in this section and all the constraints for the flow dynamics and volume dynamics are formulated. An urban traffic signal control (UTSC) problem formulation for this heterogeneous system model is provided in Section III and a distributed way to solve this UTSC problem based on Lagrangian multiplier method is shown in Section IV. The simulation-based model validation as long as the simulation results from both numerical analyses in MATLAB [27] with Gurobi [28] and simulations in VISSIM are provided in Section V. Some conclusions are drawn in Section VI.

II. MODELLING OF TRAFFIC NETWORK WITH SIGNALIZED AND NON-SIGNALIZED INTERSECTIONS

An urban traffic network consists of a set of road links connecting with each other via intersections. Each intersection consists of a number of approaches and the crossing area. An approach may have one or more lanes but has a unique, independent queue. Approaches are used by corresponding traffic streams. Two compatible streams can safely cross the intersection simultaneously, while antagonistic streams cannot. In traditional traffic signal control, a signal cycle is one repetition of the basic series of stages at an intersection, where each stage consists of simultaneous traffic light signals allowing a predefined compatible traffic streams to cross the intersection simultaneously. The duration of a cycle is called cycle time. For safety reasons, constant lost (or intergreen) times of a few seconds are necessarily inserted between consecutive stages to avoid interference between antagonistic streams. For each traffic light, the ratio of the green and red times within one cycle is called split, and a delay between the starting times of green periods of two neighboring traffic lights along the same traffic route is called offset.

A. A traffic network model

A traffic network consists of a set of road links and intersections, i.e., both signalized intersections and non-signalized intersections. Fig. 1 depicts a simple unidirectional traffic network, where each intersection has only two antagonistic traffic streams (or flows). We adopt a discrete-time model similar to a cell transmission model [19]. Let Δ be a given sampling interval, e.g., Δ = 5s. We use \( C_i(k) \) to denote the integer number of vehicles (or the volume) in the interval \( k \). Each link is partitioned into segments, where each vehicle can (approximately) traverse a segment in \( Δ \) with a free flow speed.

Let \( L \) be the set of all such one-way segments obtained above, and \( J \) the set of all (real) intersections. For each pair of consecutive segments belong to the same link, we consider them to be connected by a virtual intersection, where the corresponding traffic light is always GREEN. We denote the set of all those virtual intersections as \( J_v \) and let \( J := J_f \cup J_v \). With a slight abuse of terminology, from now on we still call each segment a link. Two groups of intersections are involved in the system, i.e., the signalized intersections and non-signalized intersections. We use \( J_s \) and \( J_n \) to denote the signalized and non-signalized ones, respectively. For each signalized intersection \( J \in J_s \), let \( \Omega_j \) be the set of stages in the intersection \( J \), and \( F_J \subseteq L \times L \) be the set of all streams in the intersection \( J \) i.e., \( (i,j) \in F_J \) means that there exists a traffic stream from the link \( i \) to the link \( j \) via the intersection \( J \). Let \( h_j : \Omega_j \rightarrow 2^{F_j} \) be the association of each stage to relevant compatible streams. Here, for simplicity we assume that for
any two different stages \( w_i \) and \( w_j \), \( h_j(w_i) \cap h_j(w_j) = \emptyset \). For the non-signalized intersection \( J \in \mathcal{J}_n \), \( F_J \subseteq \mathcal{L} \times \mathcal{L} \) is the set of all streams in the intersection \( J \) and there is no stage for such kind of intersections. For the traffic dynamics in the urban traffic system, we make the following assumption, which is suitable for a deterministic analysis.

- **A1**: The traffic demand (i.e., vehicle entrance and exit) models are known, and there is an infinite queue for each entrance demand.
- **A2**: The link turning ratios in the network are known.
- **A3**: Each vehicle inside the network will leave the network, delayed only by traffic signals.

Assumption A1 specifies that the number of vehicles requiring to enter or leave the network at relevant locations in the network is known, mainly due to historical data adjusted by real-time model identification, which is out of the scope of this paper. Once an entrance appears, it will not disappear until it is served, i.e., there is an infinite queue at each demand location to hold all vehicles, which require to enter the network. Assumption A2 can be ensured by using historical data adjusted by real-time data. Assumption A3 simply says that no vehicle will stop in the network unnecessarily, i.e., whenever allowed by the traffic signals, it will move forward. This assumption is reasonable for all normal traffic situations. In the following sections, we provide the models for the links, signalized intersections and non-signalized intersections, respectively.

### B. Link dynamics constraints

Due to the conservation of vehicles, each link \( j \in \mathcal{L} \) has the following volume dynamics:

\[
(\forall k \in \mathbb{N}) C_j(k + 1) = C_j(k) + d_j(k) - s_j(k),
\]

where \( d_j(k) \) and \( s_j(k) \) are the number of entrance vehicles and exit vehicles of the link \( j \) at \( k \) respectively, i.e.,

\[
d_j(k) = \sum_{i \in \mathcal{L} \setminus \{(i,j) \in \cup_{j \in \mathcal{J}} F_J\}} f_{ij}(k)
\]

\[
s_j(k) = \sum_{i \in \mathcal{L} \setminus \{(j,i) \in \cup_{j \in \mathcal{J}} F_J\}} f_{ji}(k)
\]

For the example shown in Fig. [4] \( d_j(k) = f_{ij}(k) \) and \( s_j(k) = f_{ji}(k) \). We define the capacity of each link based on the length of link and the minimal separations between vehicles. For example, the length of the link is 200m, the average vehicle length is 5m and the minimal separation between each two vehicles is 1m, then the capacity is about \([200/(5+1)] = 33\) (veh). Denote the capacity for link \( j \) as \( \hat{C}_j \).

\[
(\forall k \in \mathbb{N}) C_j(k) \leq \hat{C}_j,
\]

Note that no matter the link is linked to the signalized or non-signalized intersection, the equation for the link volume dynamics will always hold. Thus we do not require to distinguish the characteristics for the intersections connecting to the link. However, the traffic flow dynamics to describe the vehicle’s approaching or departing from the signalized or non-signalized intersections are different and we propose different traffic flow dynamics for them in our model.

### C. A Model for Signalized Intersections

1) **Stage constraints**: In each given interval \( k \), there exists only one active stage for an intersection \( J \in \mathcal{J}_s \), which is captured by the following constraints:

\[
(\forall w \in \Omega_J) \theta_w(k) = 0 \Rightarrow (\forall (i, j) \in h_J(w)) f_{ij}(k) = 0
\]

\[
\sum_{w \in \Omega_J} \theta_w(k) = 1
\]

\[
(\forall w \in \Omega_J)(\forall k \in \mathbb{N}) \theta_w(k) \in \{0, 1\}
\]

where \( \theta_w(k) = 0 \) and \( \theta_w(k) = 1 \) denote the RED and GREEN traffic lights associated with the stage \( w \) respectively. Condition [4a] indicates that if the stage traffic light is RED, then all relevant shifts are zero. Condition [4b] indicates that there can be only one GREEN traffic stage at any time. For all \( J \in \mathcal{J}_s \) and all \( w \in \Omega_J \) we have \( \theta_w(k) = 1 \).

2) **Flow dynamics constraints**: For each stage \( w \in \Omega_J \) and each stream \((i,j) \in h_J(w)\), the exit shift \( f_{ij}(k) \) is determined by the current upstream link volume \( C_i(j) \), the current remaining downstream link capacity \( \hat{C}_j - C_j(k) \), where \( \hat{C}_j \) is the capacity of link \( j \), and the traffic light signals in the past \( r + 1 \) time intervals \( \theta_w(k-r), \ldots, \theta_w(k) \), i.e.,

\[
f_{ij}(k) = g_{ij}(C_i(k), \hat{C}_j - C_j(k), \theta_w(k-r), \ldots, \theta_w(k)),
\]

where \( g_{ij}(\cdot) \) is a nonlinear function. The motivation behind this model is that if the stage \( w \) has been active for the past \( r \) intervals, then the drivers intend to keep a high speed \( v_{ij}(k) \) as long as the downstream link has sufficient capacity to receive the flow, i.e., \( v_{ij}(k) = l_{ij}(k)v^*_j \), where the coefficient \( l_{ij}(k) \in [0, 1] \) is monotonically nondecreasing with respect to the length of a GREEN period, except for the last few moments (or an amber or yellow light period in the current practice) when the driver needs to slow down to a complete stop to anticipate the beginning of a RED period, and \( v^*_j \) is the maximum free flow speed.

We have also done some road tests in Singapore with speed guns and the results are shown in Fig. [2]. The data is collected at an intersection in Singapore by two speed guns. We select four time intervals during two days which include the peak hours and off-peak hours and eight groups of data are recorded. The leftmost picture shows the raw data, which includes about 30000 points. The middle picture shows the processed data with error bars and the rightmost picture shows the processed data with curve fitting. In the middle and right pictures, the x-axis shows the green time length at the intersection and the y-axis shows the average speed for vehicles passing this intersection. We can conclude that with the increasing of the length of green time at the intersection, the average exit speed is increasing and finally saturated with a certain value which is always determined by the upper speed limit.

It is clear that the average exit speed in consecutive time intervals can be approximated by a set of discrete speed levels, say \( l_{ij}^0 \geq \cdots \geq l_{ij}^m > 0 \). Due to the monotonically nondecreasing nature of \( l_{ij}(k) \) with respect to the length of a GREEN period mentioned before (except for the last GREEN interval before a RED period starts), we introduce...
the following generic way of defining the speed level \(l_{ij}(k)\) based on the expected traffic light assignment:

\[
\begin{align*}
\theta_w(k + 1) = 1 & \Rightarrow l_{ij}(k) = \sum_{p=0}^{r} \delta_{ij}^p (k) t_{ij}^p & (6a) \\
\theta_w(k + 1) = 0 & \Rightarrow l_{ij}(k) = \frac{1}{2} l_{ij}(k - 1) & (6b) \\
\sum_{p=0}^{r} \delta_{ij}^p (k) - \theta_w(k) & = 0 & (6c) \\
(\forall q : 0 \leq q \leq r) (1 - \theta_w(k - q - 1)) & = 0 & (6d) \\
\prod_{p=0}^{q} \theta_w(k - p) = 1 & \land \theta_w(k + 1) = 1 \Leftrightarrow \delta_{ij}^{r-q}(k) = 1 & (6d) \\
(\forall p : 0 \leq p \leq r) \delta_{ij} \in \{0, 1\} & (6e)
\end{align*}
\]

Condition \((6a)\) says that, as long as \(k\) is not the last GREEN interval, i.e., \(\theta_w(k + 1) = 1\), the corresponding speed level \(l_{ij}(k)\) will be determined by the subsequent conditions. If the next time interval is RED, i.e., \(\theta_w(k + 1) = 0\), as stated by Condition \((6b)\), then the speed level is one half of the speed level in \(k - 1\), assuming that the traffic speed needs to be reduced linearly from \(l_{ij}(k - 1)\) to \(0\) in \(\Delta\). Condition \((6c)\) indicates that if the traffic light of the stage \(w\) is RED, i.e., \(\theta_w(k) = 0\), then \(\sum_{p=0}^{r} \delta_{ij}^p (k) = 0\), which by Condition \((6a)\) means \(l_{ij}(k) = 0\), i.e., the speed in the stage \(w\) must be zero; if the traffic light of the stage \(w\) is GREEN, i.e., \(\theta_w(k) = 1\), then by Condition \((6a)\), \(l_{ij}(k)\) can only choose one speed level because of \(\sum_{p=0}^{r} \delta_{ij}^p (k) = 1\). Conditions \((6d)\) indicates that the actual speed level depends on the number of consecutive green light intervals from \(k\) backward in time - the larger the number of consecutive green intervals, the higher the speed level.

If \(l_{ij}(k)\) is determined, the link outgoing shift \(f_{ij}(k)\) is given as follows:

\[
f_{ij}(k) = \left[ \min \{ \gamma_{ij}(k)C_i(k), \hat{C}_j - C_j(k) + s_j(k), l_{ij}(k)v_i^*d^*\Delta \} \right]
\]

where \([\cdot]\) is the largest integer not greater than input argument, \(\gamma_{ij}(k)\) is the turning ratio of vehicles in the link \(i\) towards the link \(j\) at \(k\), which is assumed to be known in advance, \(v_i^*\) is the free speed in the link \(i\), and \(d^*\) is the maximum link density, i.e., the density of the situation where all vehicles are considered having the same standard length with the minimum separation distance. Clearly, \(\sum_{j \in \mathcal{E}(i), j \in \mathcal{F}_i} \gamma_{ij}(k) = 1\), meaning that each vehicle in the link \(i\) will move into some downstream link \(j\). Conditions \((7)\) indicate that the number of vehicles in one time interval, \(f_{ij}(k)\), is the largest integer that is upper bounded by the upstream volume \(\gamma_{ij}(k)C_i(k)\) of the link \(i\), the downstream remaining capacity plus the exiting vehicles which are current in the downstream link, i.e., \(\hat{C}_j - C_j(k) + s_j(k)\), and the expected shift attainable by discounting the maximum shift \(f_{i,max} := v_i^*d^*\Delta\) by the speed level \(l_{ij}(k)\).

3) Merging constraints for signalized intersections: Mergence of multiple flows in one link is commonly seen in a traffic network. Fig 3 shows the mergence of traffic flows from left to right horizontally and from top to right (with a left turn).

\[
\begin{array}{c}
\text{Link } f_1 \\
\hline
\text{Link } f_2 \\
\hline
\text{Link } i_j
\end{array}
\]

Fig. 3. An example for traffic flow merging.
stream of vehicles immediately approaching him from its right or offside. For each intersection $J \in \mathcal{J}_n$, let $P_J : \mathcal{F}_J \to \mathbb{N}$ be a priority map such that for any two streams $\omega_1, \omega_2 \in \mathcal{F}_J$, $P(\omega_1) < P(\omega_2)$ implies that stream $\omega_1$ has a higher priority than that of $\omega_2$. With this interpretation, we have the following rule:

\[ (\forall (i, j), (i', j) \in \mathcal{F}_J)(\forall k \in \mathbb{N}) \]
\[ P(i, j) < P(i', j) \wedge f_{ij}(k) \neq 0 \Rightarrow f_{i'j}(k) = 0, \]

(8) namely, when two streams $(i, j)$ and $(i', j)$ merge into the link $j$, if $(i, j)$ has a higher priority than $(i', j)$, then there exists a non-empty flow in $(i', j)$ only if there exists no flow in $(i, j)$, i.e., vehicles in $(i', j)$ needs to give way to vehicles in $(i, j)$. The actual flow rate in $(i', j)$ during the merging process is described as follows:

\[ f_{i'j}(k) \leq \hat{C}_j - [C_j(k) - s_j(k) + f_{ij}(k)] \]

(9) The inequality (9) means that for ant time interval $k$, the give-way flow should be no more than the difference between the capacity of the downstream link $j$ and its current link volume, which takes the current incoming flow with right-of-way $f_{ij}(k)$ into account.

D. A Model for Non-Signalized Intersections

For each non-signalized intersection, we consider only the type of all-way stop-controlled intersection (AWSC), where each approaching traffic flow is required to have a full stop before going through the intersection.

![Fig. 4. An AWSC Non-signalized Intersection](image)

We make one more assumption.

- A4: Vehicles in each link move at the free flow speed whenever the distance between a vehicle in front and itself is more than the minimum separation distance.

Recall that in the previous subsection we propose a nonlinear flow rate function describing the intersection crossing model, where each driver needs to take an anticipation of the traffic light green time assignment into account. Since within a link there is no traffic light, which is equivanent to the case that there has been an infinite number of continuous green time assignement, if we imagine there is some virtual traffic light inside each link, as shown in Figure 2, by that nonlinear flow rate model, we have that each driver intends to move with the highest allowable speed, only under the constraints of the minimum separation distance. In this sense, Assumption A4 and the nonlinear flow rate model described in the previous sub-section are consistent.

With Assumption A4, we partition each link into segments such that the length of each segment is (roughly) equal to $v^* \times \lambda$, where $\lambda \in \mathbb{R}^+$ is the chosen period for each simulation step for each link connecting a non-signalized intersection. We assume that the sampling period $\Delta$ mentioned before is a multiple of $\lambda$. The reason why we choose a smaller simulation step period $\lambda$ is to improve the accuracy of the traffic model for each non-signalized intersection. Suppose a concerned link $i$ connecting to a non-signalized intersection consists of $n \in \mathbb{N}$ segments. We label them in an ascending order starting with the segment directly connecting with the non-signalized intersection with the segment ID of $0$, thus, the segment ID for the furthest segment from the non-signalized intersection is $n - 1$. For each $j \in \{0, \cdots, n - 1\}$ the link segment volume dynamic model $C_j(t)$ is the same as expression (1) shown before. But to distinguish it from the link volume dynamics related to signalized intersections, we use $t$ as the discrete time step variable. Thus, we have

\[ (\forall t \in \mathbb{N}) C_j(t + 1) = C_j(t) + d_j(t) - s_j(t), \]

(10) where $d_j(t)$ and $s_j(t)$ are the number of entrance vehicles and exit vehicles of the segment $j$ in link $i$ at $t$, respectively. For segment $0$ we define an indicator $T_i^0(t) \in \mathbb{N}$ such that $T_i^0(0) := +\infty$, and for all $t > 0$ we have

\[ T_i^0(t+1) := \begin{cases} \min\{T_i^0(t), t + 1\} & \text{if } C_i^0(t+1) > 0 \\ +\infty & \text{otherwise} \end{cases} \]

(11) The exit vehicle flow at segment $0$ can be considered as a collection of continuous sequences, during each interval among the sequence segment volume is non-zero, separated by empty periods when there is no vehicle in the segment. For each $t \in \mathbb{N}$, there can be three possibilities: (1) $t$ is the leading interval of a continuous sequence, i.e., during $t - 1$ there is no vehicle in segment $0$. In this case $T_i^0(t) := t$, indicating when the current continuous sequence starts. (2) $t$ is part of a continuous sequence but is not the leading interval. Then clearly we have $T_i^0(t) = T_i^0(t - 1) < t$, i.e., all intervals in the same continuous sequence get the leading interval as their indicators. (3) $t$ is an empty interval. In this case we have $C_i^0(t) = 0$, and $T_i^0(t) := +\infty$.

With such an indicator function for each link $i$ connecting to a non-signalized intersection $J \in \mathcal{J}_n$, we can define a virtual traffic light over the set of all links connecting to intersection $J$. Suppose each link $i \in \mathcal{L}_J$ has a unique ID $\sigma(i) \in \mathbb{N}$. We have the following traffic light green period assignment: for all $i \in \mathcal{L}_J$ and $t \in \mathbb{N}$, $\theta_i(t) = 1$ if and only if the following condition holds:

\[ \sigma(i) = \min\{\sigma(p) | p \in \mathcal{L}_J \wedge (\forall q \in \mathcal{F}_J) T_i^q(t) \leq T_i^0(t)\} \]

(12) that is, link $i$ gets a green light if and only if either none of the leading segments connecting to $J$ have vehicles and the ID of link $i$ is the smallest one (to break the tie), or there are links whose leading segments have vehicles waiting to cross $J$, and the vehicles in link $i$ reach the stop line at an earliest time and the ID of link $i$ is the smallest among all links whose
vehicles reach the stop line at the same earliest time (to break the tie). Since the link ID is unique, there can be only one link getting the green time at each interval $t$. The difference between this intersection crossing model for a non-signalized intersection and a common FCFS model in our daily life is that in the latter case the size of each continuous sequence for segment 0 of each link is always 1, but in our model this size may be bigger than 1. The consequence is that, in our model it is likely that a vehicle in link $i$ comes earlier than a vehicle in link $i'$, but because the vehicle in $i'$ is part of a continuous sequence, whose leading interval is earlier than the arrival time of the vehicle in link $i$, the vehicle in link $i'$ actually gets the green light. In this sense we can see that our model imposes FCFS over continuous vehicle columns in relevant links instead of individual vehicles, i.e., once a column of vehicles gets the green light, it will continue having it for segment 0 of each link is always 1, but in our model this is not the case. In reality, this assumption usually does not hold. But for the scheduling purpose, this is a sufficiently representative performance index. Since $s_i(k) = \sum_{j \in L_i(i,j) \in J \cup \mathcal{F}, f_{ij}(k)}$, we have our cost function as

$$\min \sum_{i \in L} \sum_{k = 1}^{N} \left[ C_i(k) - \frac{L_i}{v_{i,\text{max}}} \sum_{k = 1}^{N} f_{ij}(k) \right].$$

B. Conversions for logic and non-linear constraints

1) Conversions for signalized intersections: In the above description we have derived a linear cost function with a set of constraints describing traffic staging (4a-4c), link volume dynamics (11) and exit shifts (6a-6e). Among these constraints, (4a), (6a), (6b) and (6d) are mixed logical constraints, which can be converted into mixed integer linear constraints by adopting a transformation strategy introduced in [31], which is shown below.

Let $M$ be chosen to be sufficiently big, e.g., $M > \max_{i \in \mathcal{L}} C_i$. Then (4a) can be converted into

$$(\forall w \in \Omega_j)(\forall (i,j) \in h_j(w)) f_{ij}(k) \leq M \theta_w(k).$$

Proposition 1: Replacing Condition (4a) with Condition (17) in the urban network traffic signal scheduling formulation leads to the same solution.

Proof: To see this conversion is valid, let $\theta_w(k) = 0$, then $f_{ij}(k) \leq 0$. But since $f_{ij}(k) \geq 0$, we have $f_{ij}(k) = 0$. If $\theta_w(k) = 1$, it is trivially true that $f_{ij}(k) \leq M$ because by (7) we have $f_{ij}(k) \leq \max_{i \in \mathcal{L}} C_i < M$.

Let $[0,r - 1] := 0, \ldots, r - 1$. Conditions (6a), (6b) and (6d) are equivalent to the following: for all $q \in [0, r]$,

$$-\left(1 - \theta_w(k - q - 1)\right) + \delta_{ij}^{r-q}(k) \leq 0$$

(18a)

$$- \theta_w(k - p) + \delta_{ij}^{0}(k) \leq 0$$

(18b)

$$- \theta_w(k - q - 1) + \sum_{p=0}^{q} \theta_w(k - p) - \delta_{ij}^{r-q}(k) \leq 0$$

(18c)
where \( MC = \max \{ C^0_i(t) \} \) and \( mC = \min \{ C^0_i(t) \} \). Note that Eqn. (22b) is redundant and can be removed from the constraint set. With \( \psi^i_1(t) \) Eqn. (11) can be replaced by the following equation.

\[
T^0_i(t+1) = M + \psi^i_1(t+1) + \psi^i_1(t+1) \times \min \{ T^0_i(t), t+1 \}
\]

which can be replaced by some inequalities and logic constraints,

\[
T^0_i(t+1) \leq M + (1-\psi^i_1(t+1)) + \psi^i_1(t+1) T^0_i(t)
\]

\[
T^0_i(t+1) \leq M + (1-\psi^i_1(t+1)) + \psi^i_1(t+1) T^0_i(t)
\]

\[
[T^0_i(t+1) \geq M(1-\psi^i_1(t+1)) + \psi^i_1(t+1) T^0_i(t)]
\]

\[
\vee [T^0_i(t+1) \geq M(1-\psi^i_1(t+1)) + \psi^i_1(t+1) T^0_i(t+1)]
\]

By introducing \( \psi^i_1(t) \) and \( \psi^i_1(t) \),

\[
[T^0_i(t+1) \geq M(1-\psi^i_1(t+1)) + \psi^i_1(t+1) T^0_i(t)]
\]

\[
\leftrightarrow \psi^i_1(t+1) = 1
\]

\[
[T^0_i(t+1) \geq M(1-\psi^i_1(t+1)) + \psi^i_1(t+1) T^0_i(t)]
\]

\[
\leftrightarrow \psi^i_1(t+1) = 1
\]

Eqn. (24c) is equivalent to the mixed integer constraint set below.

\[
-M T \psi^i_1(t+1) \leq T^0_i(t+1) - M T (1-\psi^i_1(t+1))
\]

\[
-\psi^i_1(t+1) T^0_i(t) - M T
\]

\[
-(M T + \epsilon) \psi^i_1(t+1) \leq -T^0_i(t+1) + M T (1-\psi^i_1(t+1)) + \psi^i_1(t+1) T^0_i(t) - \epsilon
\]

\[
-\psi^i_1(t+1) (1-\psi^i_1(t+1)) - m T
\]

\[
-(M T + \epsilon) \psi^i_1(t+1) \leq -T^0_i(t+1) + M T (1-\psi^i_1(t+1)) + \psi^i_1(t+1) T^0_i(t) - \epsilon
\]

\[
\psi^i_1(t+1) + \psi^i_1(t+1) \geq 1
\]

Note in this constraint set, it includes some non-linear items such as \( \psi^i_1(t+1) T^0_i(t) \) and \( \psi^i_1(t+1) T^0_i(t) \). Denote \( \psi^i_1(t) T^0_i(t) = \psi^i_1(t) \) and \( \psi^i_1(t) = \psi^i_1(t) \), they can be transformed into the following linear constraints,

\[
\psi^i_1(t+1) T^0_i(t) - \psi^i_1(t+1) \leq 0
\]

\[
T^0_i(t) + \psi^i_1(t+1) T^0_i(t) - \psi^i_1(t+1) \leq T^0_i(t)
\]

\[
\psi^i_1(t+1) T^0_i(t) - \psi^i_1(t+1) T^0_i(t) \leq -T^0_i(t)
\]

\[
\psi^i_1(t+1) - T^0_i(t) - T^0_i(t) (1-\psi^i_1(t+1)) \leq -T^0_i(t)
\]

\[
\psi^i_1(t+1) (1-\psi^i_1(t+1)) \leq 0
\]

\[
\psi^i_1(t+1) - \psi^i_1(t+1) \leq 0
\]

\[
t + \psi^i_1(t+1) - \psi^i_1(t+1) \leq 0
\]

\[
\psi^i_1(t+1) - t - \psi^i_1(t+1) \leq 0
\]

For Eqn. (12) we have,

\[
(\forall q \in F) T^0_q(t) \leq T^0_q(t)
\]

\[
[T^0_q(t) \geq T^0_q(t)] \vee [T^0_q(t) \geq T^0_q(t)] \vee \cdots
\]
By introducing some binary variables $\psi^i(t) \in \{\psi^i_0(t), \psi^i_1(t), \psi^i_2(t), \psi^i_3(t), \psi^i_4(t), \psi^i_5(t)\}$,

$$\forall q \in F, \psi^i_q(t) = 1 \leftrightarrow T^q_0(t) \geq T^q_0(t)$$  \hspace{1cm} (28)

Thus, they can be transformed into the mixed integer constraints,

$$-m_T \psi^i_q(t) \leq T^q_0(t) - T^q_0(t) - m_T$$  \hspace{1cm} (29a)

$$-(M_T - \epsilon) \psi^i_q(t) \leq -T^q_0(t) + T^q_0(t) - \epsilon$$  \hspace{1cm} (29b)

Since only one state can be true in each time interval, the summation of $\psi^i_0(t)$, $\psi^i_1(t)$, $\psi^i_2(t)$ and $\psi^i_3(t)$ is one.

$$\sum \psi^i_q(t) = 1$$  \hspace{1cm} (30)

Then the constraint to determine the virtual traffic signal is shown as follows,

$$\forall q \in F, [T^q_0(t) - T^q_0(t) = 0] \leftrightarrow [\theta_i(t) = 1]$$  \hspace{1cm} (31)

which is equivalent to the inequalities shown below,

$$-m_T \theta_i(t) \leq T^q_0(t) - T^q_0(t) - m$$  \hspace{1cm} (32a)

$$-(M_T + \epsilon) \theta_i(t) \leq -T^q_0(t) + T^q_0(t) - \epsilon$$  \hspace{1cm} (32b)

$$-m_T \theta_i(t) \leq T^q_0(t) - T^q_0(t) - m$$  \hspace{1cm} (32c)

$$-(M_T + \epsilon) \theta_i(t) \leq -T^q_0(t) + T^q_0(t) - \epsilon$$  \hspace{1cm} (32d)

For Eqn. (13) - (15), please refer to the transformations for the flow dynamic constraints in signalized intersections.

When dealing with a large traffic network, solving the above MILP problem is certainly time consuming. To ensure a real-time solution to this problem, we present a distributed matheuristic approach in our previous work [24].

IV. SIMULATION RESULTS

The numerical simulations for this proposed model and traffic signal control problem are done in MATLAB with Gurobi solver. Alternatively, some real-time traffic simulations are done in VISSIM, a microscopic traffic simulator.

A. Simulation-based Model Validation in VISSIM

We use VISSIM to build an urban traffic network and validate the proposed non-signalized intersection model.

1) The VISSIM simulation procedure: For a 4-arm non-signalized intersection, there are two groups of virtual traffic signals at each intersection, i.e., the signal group for horizontal links and the signal group for vertical links, both with bi-directional traffic flows. The VISSIM simulation procedure is shown as follows. First, we set the virtual traffic signal at the non-signalized intersection to be RED for all directions and wait for the first vehicle’s coming. We set the traffic signal to be GREEN for one signal group if we detect the vehicle’s coming from current direction. The traffic signal will be kept in the next $t_{scep}$ seconds and at the end of the GREEN signal, we will detect if the link segment related to this GREEN signal is empty. If so, we give GREEN signal to the other signal group; if not, we keep the GREEN signal for this direction until the link segment is cleared.

2) The validation of non-signalized intersection model: To validate the proposed model, we compare the results from the VISSIM simulator with the results predicted by our non-signalized intersection model. Two cases are generated for this validation, i.e., the low volume and high volume scenarios, respectively. The discrete time interval is set to be 15s and the whole simulation period is 3600s. There are 240 sampling points included in each case. For each sampling point, we obtain the incoming flow for each link connected to the non-signalized intersection and outgoing flow at this intersection in VISSIM. In the mean time, we pass the same incoming flow from VISSIM to MATLAB and calculate the outgoing flow by the proposed non-signalized intersection model. The deviations between these two groups of results are shown in Fig. 5 and Fig. 6. Fig. 5 shows the deviations in the low traffic volume case and Fig. 6 shows the deviations in the high traffic volume case. We can see that the maximal deviation in this case is no more than 25% and about 80% of data are under 5% deviation. 97.92% of data are under 10% deviation, which shows the effectiveness of the proposed non-signalized flow dynamics model. Fig. 6 shows the deviations in the high traffic volume case. In this case, there does exist some data with
very large deviations because of the approximation nature of our proposed model, as mentioned in Sub-section II.D, where we impose FCFS over columns of vehicles instead of each single one. The figure shows that about 50% of the data are under 10% deviation and about 85% of the data are under 30% deviation. As a conclusion, the proposed non-signalized intersection model can provide a relatively precise description for the flow dynamics with low traffic volumes. How to develop a more accurate flow dynamic model for a non-signalized intersection with high traffic volumes remains an interesting research topic, which hopefully could be addressed in our further work.

B. Simplified Traffic Network Profile

Fig 7 shows four stages of a signalized intersection, which is currently utilized in the urban traffic network of Singapore. Different stages include a set of different traffic flows. In the fixed cycle traffic scheduling strategy, the four stages are repeated based on the order 1-2-3-4. In our model, each stage is assigned with a period in a real-time manner by the network controller with no cycle based reappearance. We use four integers, “1”, “2”, “3” and “4”, i.e., \( w \in \Omega_J = \{1, 2, 3, 4\} \), to represent four stages of one intersection shown in Fig 7.

Fig. 7. 4-stage system

For this bi-directional four-stage system, each link contains 3 traffic flows, i.e., straight-forward, left-turning and right-turning, thus each intersection contains a total of 12 traffic flows. For the intersection \( J \in \mathcal{J} \) which connected with links \( i, i_2, j_1 \) and \( j_2 \), the flows are denoted as \( f_{ij} \), where \( i \in \{i_1, i_2, j_1, j_2\} \) and \( j \in \{i_1, i_2, j_1, j_2\} \setminus \{i\} \), as shown in Fig. 8.

C. A Comparison of Fully Controlled Traffic Network, Partially Controlled Traffic Network and Fully Uncontrolled Traffic Network

We introduce a numerical comparison generated in MATLAB on the system performance among the fully controlled traffic network, partially controlled traffic network and fully uncontrolled traffic network, i.e., a traffic network with all intersections signalized, part of the intersections signalized and all intersections non-signalized. Nine cases with different traffic densities are created in this case study as shown in Table I. For the Low Traffic cases, we mean that the traffic volumes for each direction are lower than the boundaries of the gap-acceptance model; for the Medium Traffic case, the traffic volumes on some links may be higher than the gap-acceptance model and for the High Traffic case, all the traffic volumes are out of the boundaries of the gap-acceptance model, which may lead to high congestions at some intersections. All these cases are tested based on a 4-to-4 traffic grid and a total of 16 intersections are involved in this system. For the fully controlled cases, we put 16 traffic signals in this system and for the partially controlled cases 9 of the 16 intersections are signalized. The sampling time is set to be 15s and the total simulation time period is 180s. The network-wise time delays are calculated and shown in Table II.

| Case No. | Total vehicles | Total delays (sec) | Average delays (sec) |
|---------|---------------|--------------------|----------------------|
| Case 1  | 320           | 0                  | 0                    |
| Case 2  | 320           | 4860               | 486                  |
| Case 3  | 320           | 8640               | 864                  |
| Case 4  | 350           | 11700              | 1170                 |
| Case 5  | 350           | 11520              | 1152                 |
| Case 6  | 350           | 9360               | 936                  |
| Case 7  | 800           | 17640              | 1764                 |
| Case 8  | 800           | 15360              | 1536                 |
| Case 9  | 800           | 13200              | 1320                 |
Fig. 9. Number of vehicles waiting at the intersections in the 9 case studies.

For Case 1, we assume that the vehicles on all the links will not touch the threshold for the gap-acceptance model, which means that all the vehicles from different directions can pass the non-signalized intersections within one time interval unless the downstream links are blocked. Thus the network-wise traffic delay for Case 1 is 0. For Case 2, we put 9 signalized intersections. It means that for those 9 intersections only one stage can be activated for each time interval, which will block the traffic from the other direction. In this case, the network-wise traffic delay is 4860s. If we signalize all the intersections as shown in Case 3, the traffic delay will keep increasing because the freedom of flow transmission is reduced with the increasing of traffic signals. As a conclusion, for the low traffic cases, traffic signals may introduce unnecessary traffic delays to the whole system. For the Medium Traffic cases and High Traffic cases, the introducing of traffic signals does reduce the traffic delays in the system. In some scenarios such as Case 5, the signalized intersections do not have much more improvements. However, when the traffic congestions are heavy, the signalized intersections will help to significantly reduce the delay time, as shown in Case 8 and Case 9.

In Fig. 9 we demonstrate the vehicles delayed at the signalized or non-signalized intersections. The circles denote the signalized intersections and the rectangles denote the non-signalized intersections. The color changes from dark blue to dark brown, which denotes the level of congestions from low to high. Numbers in the circles or rectangles are the number of vehicles delayed at the intersections from all the directions.

D. VISSIM Simulation in Singapore Traffic Networks

A simulation platform is developed in VISSIM based on Jurong West Area in Singapore which is shown in Fig. 10, which consists of 9 intersections and 40 links. The VISSIM simulation platform is connected to MATLAB via COM interface and the traffic signal controller is developed in MATLAB with Gurobi optimization solver. We test this case study with fully uncontrollable network, partially controllable network and fully controllable network with low traffic, medium traffic and high traffic volumes as well.

V. CONCLUSIONS

In this paper, a novel model for the urban traffic system is proposed to describe the heterogeneous traffic system with signalized intersections and non-signalized intersections. The proposed model is validated based on the simulations in VISSIM. Moreover, an urban traffic signal control problem formulation for the traffic network with signalized and non-signalized intersections is proposed as a mixed integer programming problem. Some comparisons among the fully controlled, partially controlled and fully uncontrollable traffic system are provided based on the numerical analyses and we show the potential applications in the traffic system design, which leaves the door open for developing a systematic planning approach on deciding what traffic junctions require signal control to ensure a good traffic control performance, thus, have a great social and economic potentials, considering that it is rather expensive to have signal control in an urban area.
REFERENCES

[1] M. S. Shirazi, and B. T. Morris. Looking at Intersections: A Survey of Intersection Monitoring, Behavior and Safety Analysis of Recent Studies. IEEE Transactions on Intelligent Transportation Systems, 18(1), pp.42-24, 2017.

[2] D. I. Robertson. TRANSYT method for area traffic control. Traffic Engineering & Control, vol. 10, pp. 276-281, 1969.

[3] H. J. Ruskin, and R. Wang. Modeling traffic flow at an urban unsignalized intersection. In International Conference on Computational Science, pp. 381-390, Springer Berlin Heidelberg, 2002.

[4] E. Chevalier, and L. Leclercq. A macroscopic theory for unsignalized intersections. Transportation Research Part B: Methodological, 41(10), pp.1139-1150, 2007.

[5] W. Brilon, and N. Wu. Capacity at unsignalized intersections derived by conflict technique. Transportation Research Record: Journal of the Transportation Research Board, (1776), pp.82-90, 2001.

[6] N. Wu. A universal procedure for capacity determination at unsignalized (priority-controlled) intersections. Transportation Research Part B: Methodological, 35(6), pp.593-623, 2001.

[7] N. Wu. A new model for estimating critical gap and its distribution at unsignalized intersections based on the equilibrium of probabilities. In Proceeding of the 5th International Symposium on Highway Capacity and Quality of Service, pp. 1-10, 2006, July.

[8] M. E. Foulaadvand, and S. Bellbasi. Vehicular traffic flow at a non-signaled intersection. Journal of Physics A: Mathematical and Theoretical, 40(29), p.8289, 2007.

[9] Q. S. Wu, X. B. Li, M. B. Hu, and R. Jiang. Study of traffic flow at an unsignalized T-shaped intersection by cellular automata model. The European Physical Journal B-Condensed Matter and Complex Systems, 39(2), pp.265-269, 2005.

[10] R. E. Allsop. SIGSET: A computer program for calculating traffic capacity of signal-controlled road junctions. Traffic Engineering & Control, vol. 12, pp. 58-60, 1971.

[11] R. E. Allsop. SIGCAP: A computer program for assessing the traffic capacity of signal-controlled road junctions. Traffic Engineering & Control, vol. 17, pp. 338-341, 1976.

[12] M. T. Li, & A. C. Gan. Signal timing optimization for oversaturated networks using TRANSYT-7E. In Proceeding of the 78th Annual Meeting Transportation Research Board, 1999.

[13] P. B. Hunt, D. L. Robertson, & R. D. Bretherton. The SCOOT on-line traffic signal optimization technique. Traffic Engineering & Control, vol. 23, pp. 190-192, 1982.

[14] D. Robertson, & R. Bretherton. "Optimizing networks of traffic signals in real time: The SCOOT method." IEEE Transactions on Vehicular Technology, vol. 40 no. 1, 1991.

[15] N. H. Gartner. OPAC: A demand-responsive strategy for traffic signal control. U.S. Department of Transportation, Washington, DC, Transportation Research Record 906, 1983.

[16] J. L. Farges, J. J. Henry, & J. Tufal, THE PRODYK real-time traffic algorithm. In Proceeding of 4th IFAC Symposium on Transportation Systems, pp. 307-312, 1983.

[17] M. J. Lighthill, & G. B. Whitham. On kinematic waves. II. A theory of traffic flow on long crowded roads. In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 229, no. 1178, pp. 317-345, 1955.

[18] P. I. Richards. Shock waves on the highway. Operations research, vol. 4, no. 1, pp.42-51, 1956.

[19] C. F. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B: Methodological, vol. 28, no. 4, pp. 269-287, 1994.

[20] C. F. Daganzo. A finite difference approximation of the kinematic wave model of traffic flow. Transportation Research Part B: Methodological, vol. 29, no. 4, pp. 261-276, 1995.

[21] H. K. Lo. "A novel traffic signal control formulation." Transportation Research Part A: Policy and Practice 33, no. 6 (1999): 433-448.

[22] H. K. Lo. "A cell-based traffic control formulation: strategies and benefits of dynamic timing plans." Transportation Science 35, no. 2 (2001): 148-164.

[23] W. H. Lin, & C. H. Wang. "An enhanced 0-1 mixed-integer LP formulation for traffic signal control." Intelligent Transportation Systems, IEEE Transactions on 5, no. 4 (2004): 238-245.

[24] Y. Zhang, R. Su, K. Gao. Urban road traffic light real-time scheduling. In proceeding of the 54th IEEE Conference on Decision and Control (CDC), Osaka, Osaka, 2015.

Yicheng Zhang (S’15) Yicheng Zhang received his Master degree from University of Science and Technology of China in 2014. He is currently a PhD candidate in School of Electrical and Electronic Engineering at Nanyang Technological University, Singapore. His current research interests include optimization model, method-based fault diagnosis, smart grid and intelligent transportation systems.

Rong Su (M’11-SM’14) Dr Su Rong obtained his Bachelor of Engineering degree from University of Science and Technology of China, and Master of Applied Science degree and PhD degree from University of Toronto. Since then he was affiliated with University of Waterloo and Technical University of Eindhoven before he joined Nanyang Technological University in 2010. Dr Su’s research interests include discrete-event system theory, model-based fault diagnosis, operation planning and scheduling and multi-agent systems with applications in flexible manufacturing, intelligent transportation, human-robot interface, power management and green building. So far he has been involved in several projects sponsored by Singapore National Research Foundation ( NRF), Singapore Agency of Science, Technology and Research (A*STAR), Singapore Ministry of Education (MoE), Singapore Civil Aviation Authority ( CAAS) and Singapore Economic Development Board (EDB). Dr Su is a senior member of IEEE and an associate editor for Transactions of the Institute of Measurement and Control and Journal of Control and Decision. He is also the Chair of the Technical Committee on Smart Cities in the IEEE Control Systems Society.
Yi Zhang (S’17) Yi Zhang received her Bachelor degree from China in Shandong University in 2014. She is currently a PhD candidate at Nanyang Technological University, Singapore. Her current research interests include model-based traffic signal scheduling and multi-directional pedestrian-flow simulation.

Chunyang Sun Chunyang Sun received B.Eng. (EEE) from Nanyang Technological University (NTU), Singapore, 2016. He was a software engineer in Innogrity Pte Ltd (Singapore) from Sep. 2012 to Aug. 2014. Since January 2016, he was a project officer in the School of Electrical and Electronic Engineering (EEE), NTU, Singapore. His research interests include traffic simulation and software development.