

An integrated control model for freeway corridor under non-recurrent congestion

Yue Liu^{*}; Gang-Len Chang, Jie Yu

Abstract-This study presents an integrated model and its solution algorithm for freeway corridor control during incident management. With a parallel arterial as the detour route, the proposed model aims to produce the optimal diversion rates from the freeway mainline to relieve the congestion at the incident segment, and concurrently adjust signal timings at the arterial intersections to best accommodate the detour traffic. Different from previous studies, the presented model and algorithm have the following two critical features: (1) modeling explicitly the evolution of detour traffic along the ramps and surface streets with a set of dynamic network flow formulations so as to capture the local bottlenecks caused by demand surge due to diversion operations, and to properly set the responsive signal timing plans; and (2) developing a multi-objective optimization framework to maximize the utilization of the available corridor capacity via detour operations, but not to incur excessive congestion on the arterials and ramps. This study employs a Genetic Algorithm (GA) based heuristic to efficiently yield the reliable solution depending on decision maker's preference. Extensive numerical tests on a segment along the I-95 corridor with its neighboring arterials have demonstrated the potential of the developed model for integrated freeway corridor control.

Index Terms: Freeway operation; Traffic flow model; Diversion; Integrated traffic control

I. INTRODUCTION

TRAFFIC delays on urban freeways due to congestion have significantly undermined the efficiency and mobility of the highway systems in the United States. Most of those delays are due to non-recurrent traffic congestion caused by the growing demand and increasing accidents on critical metropolitan corridors. As reported in the literature, non-recurrent traffic congestion contributes up to 60 percent of the total freeway delays in the United States. In most scenarios, if implementing the routing and control strategies in a timely manner, motorists can circumvent the congested segments by detouring through parallel arterials. To contend with this vital operational

issue, transportation professionals have proposed various types of optimal control models on each individual component of the freeway corridor including: ramp metering [1-12], freeway control [13-17], diversion control, and route guidance [18-24]. Certainly, those research efforts have advanced the development of control strategies and operational guidelines for freeway congestion management. However, the lack of coordinated operations among those control modules have resulted in underuse of the corridor capacity and incurred excessive delays.

Most of existing studies on this subject focus on developing non-linear models for simultaneously or sequentially optimizing various corridor control strategies. For example, Cremer and Schoof [25] first formulated an integrated control model with a two-level optimization framework, where its upper-level is for diversion optimization and the lower-level for optimization of ramp metering, speed limit, and intersection signal timings. Their proposed model does not optimize all control variables concurrently, and not include the coordination of signals on surface streets. Zhang and Hobeika [26] proposed a nonlinear programming model to determine the diversion routes and rates, ramp metering rates, and arterial signal timings for a freeway corridor under incident conditions. Their optimization model is capable of preventing congestion by limiting queue lengths with constraints in the objective function, but not delays at on-ramps and off-ramps. Chang et al. [27] presented a dynamic control model for a commuting corridor, including a freeway and a parallel arterial. With the assumption that traffic diversion and route choice of all traffic demands were predictable, they included ramp metering and intersection signal timing variables in a single optimization model and solve it simultaneously in a system-optimal fashion.

To improve the computing efficiency but not compromising the accuracy, Papageorgiou [28] developed a linear optimal-control model to design integrated control strategies for traffic corridors, including both motorways and signal-controlled urban roads based on the store-and-forward modeling philosophy. Wu and Chang [29] formulated a linear programming system for integrated corridor control in which the flow-density relation was approximated with a piece-wise linear function to facilitate the use of a successive linear programming algorithm for global optimality. Van den Berg et al. [30] proposed a model predictive control approach for mixed urban and freeway networks, based on the enhanced macroscopic traffic flow models in which traffic flow evolution on ramps have been explicitly captured.

Despite the promising progress from those integrated control models, development of effective and efficient integrated control strategies for an urban freeway network under the incident situation remains a challenging task, especially regarding the following issues:

- The evolution of diversion traffic along the detour route and its impacts on intersection turning patterns have not been modeled explicitly in a dynamic context. Most previous studies address this issue either by projecting the turning proportions at intersections, based on assumed dynamic OD and travel time information, which may not be available in a real-world application, or by applying a fixed additional amount of flow to the impacted movement, which often does not reflect changes in the time-dependent pattern;
- More realistic formulations of discharge flows at intersection approaches are needed, since most studies model dynamic queue evolution either at a link-based level or at an individual-movement-based level, which could result in either difficulty in integrating with multiple signal phases or inaccuracy in modeling the queue discharging rates in a shared lane;
- Existing approaches are insufficient to address severe congestion due to non-recurrent incidents. For example, they have not adequately captured the spillback queue interactions and blocking effects among different lanes caused by the demand surges of a specific movement due to diversion;
- The coordination of arterial signal controllers has not been concurrently considered in the control process; and
- The inherently multi-objective nature of the integrated corridor control has not been fully addressed. Most previous studies proposed an optimal control model with one objective — either to minimize total network travel time or to maximize its total throughput. However, the single control objective may result in unbalanced travel times between the detour route and the freeway mainline and not fully capture the trade-offs between the target freeway and available detour routes.

To address the above critical issues, this study presents an integrated model and its solution for freeway corridor control under non-recurrent congestion. It offers an effective tool to better capture the temporal/spatial interactions of traffic over the corridor network, including the freeway segments, arterials, and ramps. The proposed new lane-group-based concept also provides a more accurate representation of the relationships between the arriving and departing flows under various intersection channelization (e.g. shared lanes) and signal control environment. By properly integrating with the traditional destination-dependent sub-flow modeling concept, the proposed model can capture explicitly the evolution of detour traffic along the ramps and surface streets as well as

the resulting local bottlenecks caused by the surge in traffic demand levels. Such a unique modeling feature can more accurately and effectively set the control variables in the overall corridor optimization process.

The remainder of this paper is organized as follows. Section II will detail the research scope and assumptions for formulating the integrated corridor control model. Section III presents the enhanced network flow equations, which include three modules for modeling traffic dynamics in the arterial, freeway section, and on-off ramps. Section IV illustrates the formulations of the optimization model for integrated control. Section V introduces the solution framework for the proposed control model. Section VI evaluates the proposed model with extensive experimental tests. Section VII concludes the work.

II. RESEARCH SCOPE AND ASSUMPTIONS

For convenience of illustration, this study focuses on a control area, which includes a segment of the freeway mainline experiencing an incident, the on-ramps and off-ramps right upstream and downstream to the incident location, and the connecting parallel arterial (see Fig. 1). Note that it can be a natural extension of the hereafter proposed model to accommodate multiple segments of a corridor network when necessary. To ensure that the proposed formulations for integrated corridor control are tractable and also realistically reflect the real-world constraints, this study has employed the following assumptions:

- Traffic is diverted to the arterial through the off-ramp just upstream to the incident section, and will be guided back to the freeway. The compliance rate for drivers is assumed to be known or obtainable from the on-line surveillance system deployed in the control area;
- Normal traffic patterns, including off-ramp exit rates, normal traffic getting into the freeway via the on-ramp, and existing arterial intersection turning proportions, are assumed to be stable and not impacted by the detour traffic, or the impact can be estimated (will be addressed by the proposed model);
- A common cycle length is assumed for all intersections in the arterial, and the phase sequence is pre-set; and
- The entire control time horizon (H) is decomposed into a series of control time intervals (T), and control decisions are optimized over each successive time interval.

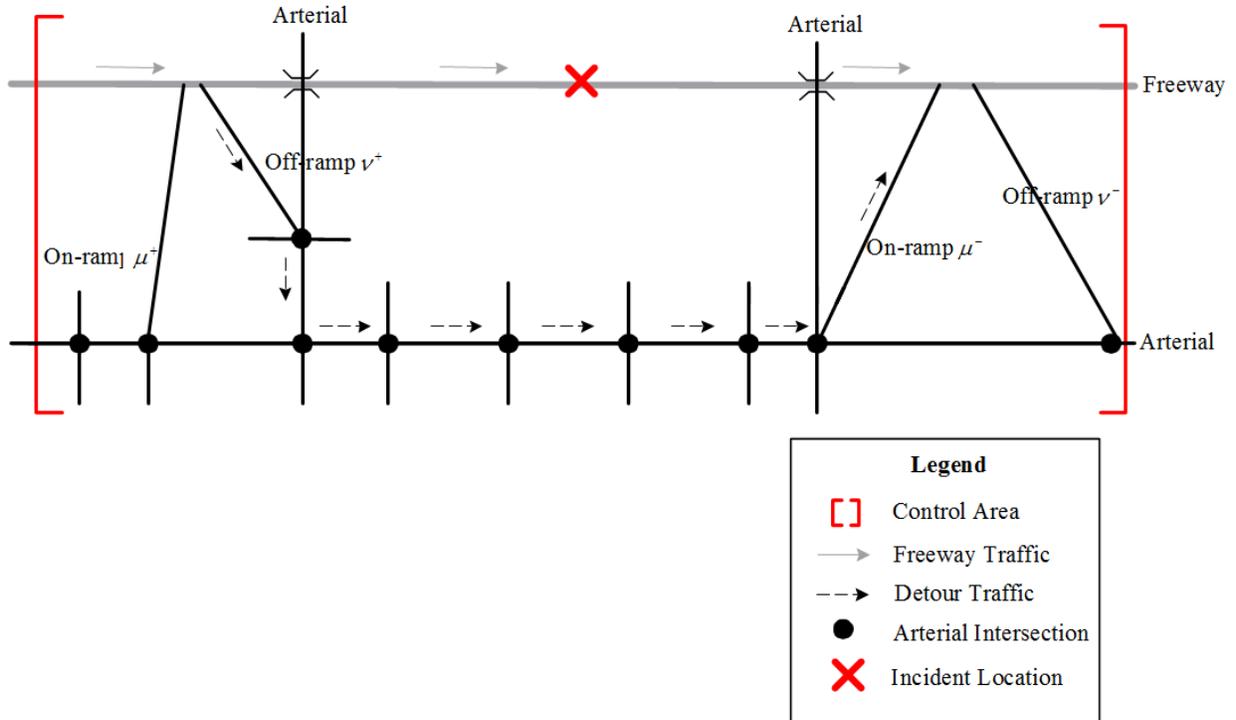


Fig. 1 Scope of the integrated control model

III. THE NETWORK FLOW FORMULATION

A. Arterial Flow Dynamics

This section illustrates an enhanced traffic flow model that can precisely project the time-varying impacts of detour traffic on the arterial flow patterns by seamlessly incorporating the sub-flow concept in the lane-group-based model [31][32]. The proposed model consists of the following six module sets: demand entries, upstream arrivals, joining the end of queue, merging into lane groups, departing process, and flow conservation. Fig. 2 shows an example of a typical traffic evolution process on an internal arterial link. The key features of the proposed model lie in its capability of

- Tracking the evolution of detour traffic along the arterial and its impact on each movement;
- Capturing the evolution of physical queues with respect to the signal status, arrivals, and departures;
- Modeling the merging and splitting of vehicle movements at intersections; and
- Capturing local bottlenecks such as overflows and blockages caused by dramatic changes in demand levels and patterns due to diversion operations.

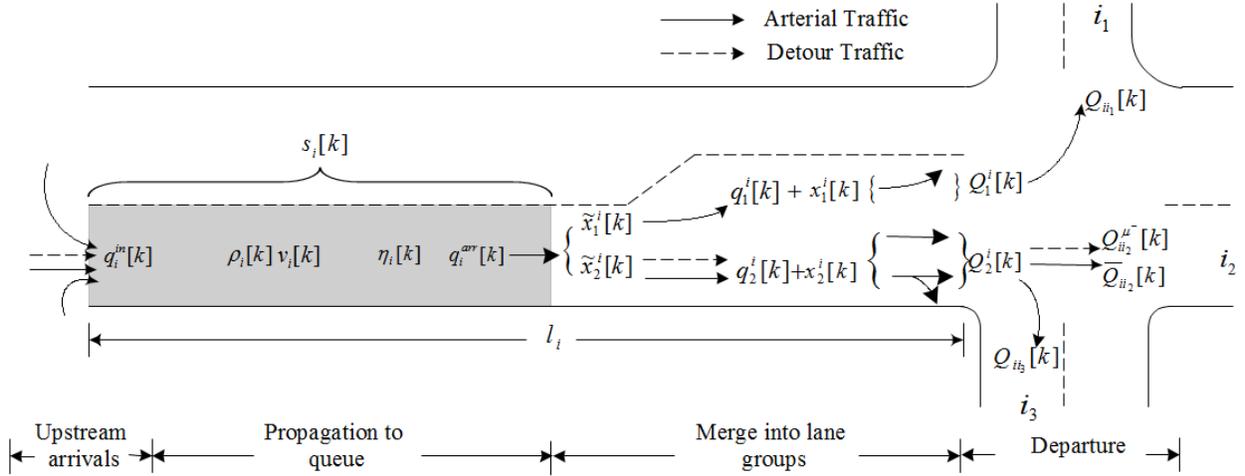


Fig. 2 Traffic flow evolution at an internal arterial link considering the detour impact

To facilitate the model presentation, the notations used hereafter are summarized in Table 1.

TABLE 1 List of key variables used in the traffic model for arterial dynamics

Δt	Update interval of arterial dynamics (in seconds)
k	Time step index corresponds to time $t = k\Delta t$
$n, n \in S_N$	Index of arterial intersections
$i, i \in S^U$	Index of links
S^{OUT}	Set of outgoing boundary links in the arterial
μ^+, ν^+	Index of the incident upstream on-ramp and off-ramp, respectively (see Fig. 1)
μ^-, ν^-	Index of the incident downstream on-ramp and off-ramp, respectively (see Fig. 1)
$p, p \in P_n$	Index of signal phase at the intersection n
S_r	Set of traffic demand entries in the arterial network
$\Gamma(i), \Gamma^{-1}(i)$	Set of upstream and downstream links of link i
l_i, n_i, N_i, Q_i	Length (in meters), # of lanes, storage capacity (in vehs), and discharge capacity (in vph) of link i
$m, m \in S_i^M$	Index of lane groups at link i
$\delta_m^{ij}, j \in \Gamma^{-1}(i)$	A binary indicating whether the movement from link i to j uses lane group m
N_m^i, Q_m^i	Storage capacity (in vehs) and discharge capacity (in vph) for lane group m
$\Omega^i[k]$	Blocking matrix between lane groups at link i
$\omega_{m'm}^i[k] \in \Omega^i[k]$	Blocking coefficient between lane group m' and m at step k
$D_r[k], r \in S_r$	Flow rate generated at demand entry r at step k (in vph)
$IN_r[k], r \in S_r$	Flow rate entering the link from demand entry r at step k (in vph)
$w_r[k], r \in S_r$	Queue waiting on the entry r at step k (in vehs)
$q_i^{in}[k]$	Num. of upstream inflow vehicles of link i at step k (in vehs)
$\bar{N}_i[k]$	Num. of vehicles from normal arterial traffic at link i at step k (in vehs)
$N_i^{\mu^-}[k]$	Num. of detour vehicles heading to downstream on-ramp μ^- at link i at step k (in vehs)
$\bar{\gamma}_{ij}[k], j \in \Gamma^{-1}(i)$	Relative turning proportion of normal arterial traffic from link i to j
$\gamma_{ij}^{\mu^-}, j \in \Gamma^{-1}(i)$	A binary value indicating whether detour traffic at link i heading to downstream on-ramp μ^- will use downstream link j or not
$\eta_i[k]$	Fraction of normal arterial traffic in total traffic at link i at step k

$s_i[k]$	Available space of link i at step k (in vehs)
$x_i[k]$	Num. of vehicles in queue at link i at step k (in vehs)
$q_i^{arr}[k]$	Num. of vehicles arriving at end of queue of link i at step k (in vehs)
$q_m^{i,pot}[k]$	Num. of vehicles potentially to merge into lane group m of link i at step k (in vehs)
$q_m^i[k]$	Num. of vehicles join the queue of lane group m at step k (in vehs)
$x_m^i[k]$	Queue length of lane group m at link i at step k (in vehs)
$\tilde{x}_m^i[k]$	Num. of arrival vehicles with destination to lane group m queued outside the approach lanes due to blockage at link i at step k (in vehs)
$\lambda_m^{ij}[k], j \in \Gamma^{-1}(i)$	Percentage of traffic in lane group m going from link i to j
$\bar{\lambda}_m^{ij}[k], j \in \Gamma^{-1}(i)$	Percentage of normal arterial traffic in lane group m going from link i to j
$Q_m^i[k]$	Num. of vehicles depart from lane group m at link i at step k (in vehs)
$Q_{ij}^{pot}[k]$	Num. of vehicles potentially depart from link i to link j at step k (in vehs)
$Q_{ij}[k]$	Total flows actually depart from link i to link j at step k (in vehs)
$\bar{Q}_{ij}[k]$	Normal arterial traffic flows actually depart from link i to link j at step k (in vehs)
$Q_{ij}^{\mu^-}[k]$	Detour traffic flows heading to downstream on-ramp μ^- actually depart from link i to link j at step k (in vehs)
$g_n^p[k]$	Binary value indicating whether signal phase p of intersection n is green or not at step k

Demand entries

Arterial demand entries are modeled as follows:

$$IN_r[k] = \min \left[D_r[k] + \frac{w_r[k]}{\Delta t}, Q_i, \frac{s_i[k]}{\Delta t} \right] \quad (1)$$

$$w_r[k+1] = w_r[k] + \Delta t [D_r[k] - IN_r[k]] \quad (2)$$

Equation (1) indicates that the flow enters downstream link i from demand entry r depends on the existing flows queuing at r , discharge capacity of the link i , and the available space in the link i . Equation (2) updates the queue waiting at the demand entry during each time step.

Upstream arrivals

Upstream arrivals are formulated to depict the evolution of flows arriving to the upstream of the link over time.

Eqns. (3) and (4) define the flow dynamics for different types of links.

For internal arterial links, inflows to link i can be formulated as the sum of actual departure flows from all upstream links, including both normal arterial traffic and detour traffic:

$$q_i^{in}[k] = \sum_{j \in \Gamma(i)} \bar{Q}_{ji}[k] + \sum_{j \in \Gamma(i)} Q_{ji}^{\mu^-}[k] \quad (3)$$

For source links (connected with demand entry r), inflows can be formulated as:

$$q_i^{in}[k] = IN_r[k] \cdot \Delta t \quad (4)$$

Where, $\bar{Q}_j[k]$ and $Q_j^{\mu^-}[k]$ represent the actual flows departing from upstream link j to link i for normal arterial traffic and detour traffic, respectively.

Joining the end of queue

This module represents the evolution of upstream inflows to the end of queue with the average approaching speed. The mean speed of vehicles, $v_i[k]$, depending on the density of the segment between the link upstream and the end of queue, $\rho_i[k]$, can be described with the following equation [33]:

$$v_i[k] = \begin{cases} v_i^{free}, & \text{if } \rho_i[k] < \rho^{\min} \\ v^{\min} + (v_i^{free} - v^{\min}) \cdot \left[1 - \left(\frac{\rho_i[k] - \rho^{\min}}{\rho^{jam} - \rho^{\min}} \right)^\alpha \right]^\beta, & \text{if } \rho_i[k] \in [\rho^{\min}, \rho^{jam}] \\ v^{\min}, & \text{if } \rho_i[k] > \rho^{jam} \end{cases} \quad (5)$$

Where, $v_i[k]$ represents the mean approaching speed of vehicles from upstream to the end of queue at link i at step k ; ρ^{\min} is the minimum critical density below which traffic at link i moves at the free flow speed (v_i^{free}); v^{\min} is the minimum traffic flow speed corresponding to the jam density (ρ^{jam}); and α, β are constant model parameters to be calibrated. The density of the segment from link upstream to the end of queue, $\rho_i[k]$, is computed with the following equation:

$$\rho_i[k] = \frac{\bar{N}_i[k] + N_i^{\mu^-}[k] - x_i[k]}{n_i \left(l_i - \frac{1000 \cdot x_i[k]}{n_i \cdot \rho^{jam}} \right)} \quad (6)$$

Where, $\bar{N}_i[k] + N_i^{\mu^-}[k] - x_i[k]$ represents the number of vehicles (both normal arterial traffic and detour traffic) moving at the segment between the link upstream and the end of queue, and $l_i - 1000 \cdot x_i[k] / (n_i \cdot \rho^{jam})$ depicts the length of that segment over time. Then, the number of vehicles arriving at the end of queue at link i can be dynamically updated with:

$$q_i^{arr}[k] = \min \left\{ \rho_i[k] \cdot v_i[k] \cdot n_i \cdot \Delta t, \bar{N}_i[k] + N_i^{\mu^-}[k] - x_i[k] \right\} \quad (7)$$

Where, $\rho_i[k] \cdot v_i[k] \cdot n_i \cdot \Delta t$ represents the flows potentially arriving at the end of queue at time step k , which is limited by $\bar{N}_i[k] + N_i^{\mu^-}[k] - x_i[k]$.

Merging into lane groups

After vehicles arrive at the end of a link queue, they will try to change lanes and merge into different lane groups based on their destinations. Most previous studies assume that the arriving vehicles could always merge into their destination lanes without being blocked. However, such an assumption may not be realistic under the following scenarios: (I) the intended lane group has no more space to accommodate arriving vehicles (e.g., a fully occupied left-turn bay); and (II) the overflowed queues from other lane groups block the target lane group (see Fig. 3). Therefore, arriving vehicles that could not merge into their destination lane group m due to either overflows or blockage will form spillback to neighboring lanes, denoted by $\tilde{x}_m^i[k]$. Due to the detour operation, the demand level for a target arterial movement could surge to the bottleneck level, as mentioned in Scenario I and Scenario II. Thus, it is critical for the proposed model to capture the traffic interactions on these bottlenecks and to reflect their impacts on the control parameter optimization.

To illustrate such scenarios, it should be noted that the number of vehicles allowed to merge into lane group m at time step k depends on the available storage capacity of the lane group, given by:

$$\max\{N_m^i - x_m^i[k], 0\} \quad (8)$$

Further, the aforementioned blocking impacts between different lane groups can be classified as complete blockage and partial blockage (shown in Fig. 3). In order to model such queue interactions between every pair of lane groups in a dynamic manner, we define a blocking matrix for each arterial link i , denoted by $\Omega^i[k]$. The dimension of the blocking matrix is $M_i \times M_i$, and M_i is the number of lane groups at link i . The matrix element, $\omega_{m'm}^i[k]$, takes a value between 0 and 1 to depict the blocking effect on lane group m due to the queue spillback at lane group m' at time step k . In this paper, we modeled $\omega_{m'm}^i[k]$ as follows:

$$\omega_{m'm}^i[k] = \begin{cases} 1 & x_{m'}^i[k] > N_{m'}^i, \text{ complete blockage} \\ \phi_{m'm} \cdot \frac{q_{m'}^{i,pot}[k]}{\sum_{m \in \mathcal{S}_i^M} q_m^{i,pot}[k]} & x_{m'}^i[k] > N_{m'}^i, \text{ partial blockage} \\ 0 & \text{no blockage or } x_{m'}^i[k] \leq N_{m'}^i \end{cases} \quad (9a)$$

$$q_m^{i,pot}[k] = \tilde{x}_m^i[k] + \sum_{j \in \Gamma^{-1}(i)} q_j^{arr}[k] \cdot [\eta_i[k] \bar{\gamma}_{ij}^-[k] + (1 - \eta_i[k]) \gamma_{ij}^{\mu^-}[k]] \cdot \delta_m^{ij} \quad (9b)$$

Where, $q_i^{arr}[k]$ is the total flows arriving at the end of queue of link i at time step k ; $q_i^{arr}[k] \cdot \eta_i[k] \cdot \bar{\gamma}_{ij}[k]$ represents the normal arterial traffic flow going to link j at time step k , and $q_i^{arr}[k] \cdot (1 - \eta_i[k]) \cdot \gamma_{ij}^{\mu^-}$ denotes the detour traffic flow going to link j at time step k ; δ_m^{ij} is a binary value indicating whether traffic going from link i to j uses lane group m . Hence, one can approximate $\tilde{x}_m^i[k] + \sum_{j \in \Gamma^{-1}(i)} q_i^{arr}[k] \cdot [\eta_i[k] \bar{\gamma}_{ij}[k] + (1 - \eta_i[k]) \gamma_{ij}^{\mu^-}] \cdot \delta_m^{ij}$ as the potential level of flows that may merge into lane group m at time step k , denoted as $q_m^{i,pot}[k]$.

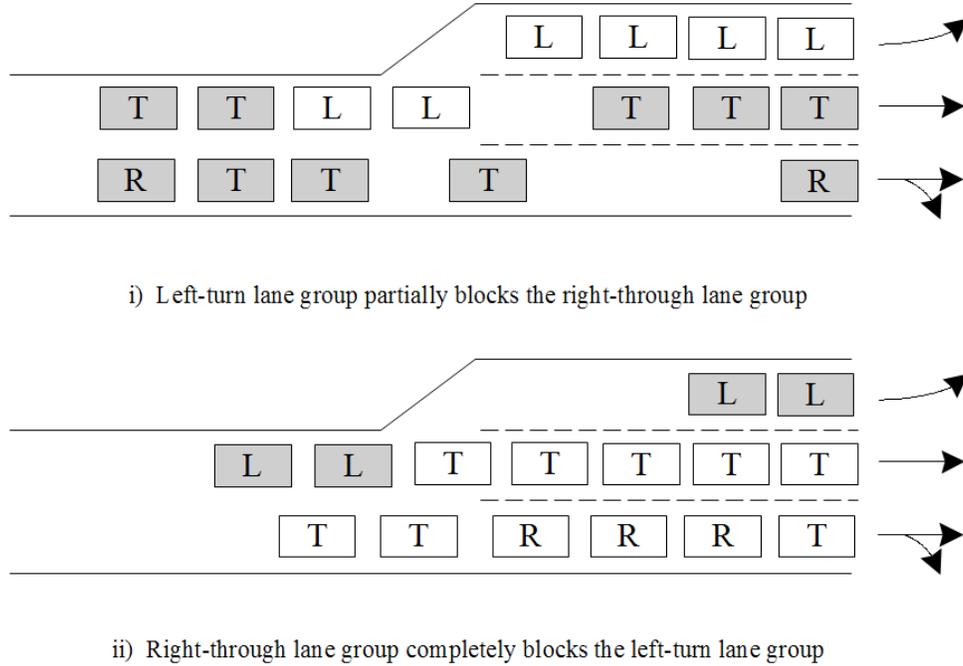


Fig. 3 Blockages between Lane Groups

To ensure the blocking matrix effectively discriminate complete blockage from partial blockage, one can pre-specify the blocking impact between any given pair of lane groups based on the geometric features in a target intersection approach. For example in Fig. 3, the impact of the left-turn lane group on the right-through lane group is observable to be a partial blockage, while the impact of the right-through lane group on the left-turn lane group is a typical complete blockage. Thus, at each time step, the model shall be able to “understand” the blocking types and evaluate each element in the blocking matrix if any queue spillback among the target lane groups has occurred.

As shown in Equation 9(a), for the complete blockage or no blockage cases, one can set $\omega_{m'm}^i[k]$ to be 1 or 0, based on the geometric features of the approach (shown in Fig. 3-ii). For the partial blockage case, $\omega_{m'm}^i[k]$ can be approximated by $\phi_{m'm} \cdot q_m^{i,pot}[k] / \sum_{m \in S_i^M} q_m^{i,pot}[k]$; where, $\phi_{m'm}$ is a constant parameter between 0 and 1 and is related to driver's response to the lane blockage and geometry features; and

$q_m^{i,pot}[k] / \sum_{m \in S_i^M} q_m^{i,pot}[k]$ approximates the fraction of merging lanes occupied by the overflowed traffic from lane group m' at time step k . Taking the link shown in Fig. 3 as an example, there are two lane groups in the link: left-turn and right-through (named as L and R-T, respectively). Therefore, the blocking matrix is in 2×2 dimension,

constructed as $\begin{bmatrix} \omega_{L,L}^i[k] & \omega_{L,R-T}^i[k] \\ \omega_{R-T,L}^i[k] & \omega_{R-T,R-T}^i[k] \end{bmatrix}$. The elements in the matrix will be updated as follows:

$$\omega_{L,L}^i[k] = \begin{cases} 1 & \text{if } x_L^i[k] > N_L^i \\ 0 & \text{if } x_L^i[k] \leq N_L^i \end{cases}, \text{ L blocks itself;}$$

$$\omega_{L,R-T}^i[k] = \begin{cases} \phi_{L,R-T} \cdot \frac{q_L^{i,pot}[k]}{q_L^{i,pot}[k] + q_{R-T}^{i,pot}[k]}, & \text{if } x_L^i[k] > N_L^i \\ 0 & \text{if } x_L^i[k] \leq N_L^i \end{cases}, \text{ L partially blocks R-T;}$$

$$\omega_{R-T,L}^i[k] = \begin{cases} 1 & \text{if } x_{R-T}^i[k] > N_{R-T}^i \\ 0 & \text{if } x_{R-T}^i[k] \leq N_{R-T}^i \end{cases}, \text{ R-T completely blocks L;}$$

$$\omega_{R-T,R-T}^i[k] = \begin{cases} 1 & \text{if } x_{R-T}^i[k] > N_{R-T}^i \\ 0 & \text{if } x_{R-T}^i[k] \leq N_{R-T}^i \end{cases}, \text{ R-T blocks itself;}$$

Considering the impact of blocking matrix, the number of vehicles allowed to merge into lane group m at time step k is restricted by:

$$\max \left\{ q_m^{i,pot}[k] \cdot \left[1 - \sum_{m' \in S_i^M} \omega_{m'm}^i[k] \right], 0 \right\} \quad (10)$$

Where, according to the definition of $\omega_{m'm}^i[k]$, $1 - \sum_{m' \in S_i^M \wedge m' \neq m} \omega_{m'm}^i[k]$ is the residual fraction of capacity to

accommodate those potential merging vehicles to lane group m .

Finally, the number of vehicles that are allowed to merge into lane group m at time step k should be the minimum value of Eqns. (8) and (10), given by:

$$q_m^i[k] = \min \left\{ \max \{ N_m^i - x_m^i[k], 0 \}, \max \left\{ q_m^{i,pot}[k] \cdot \left[1 - \sum_{m' \in S_i^M} \omega_{m'm}^i[k] \right], 0 \right\} \right\} \quad (11)$$

Departing process

The number of vehicles potentially departing from link i to link j at time step k is given by:

$$Q_{ij}^{pot}[k] = \sum_{m \in S_i^M} \min \{ q_m^i[k] + x_m^i[k], Q_m^i \cdot g_n^p[k] \} \cdot \lambda_m^{ij}[k] \quad (12)$$

And, $\lambda_m^{ij}[k]$ can be estimated by:

$$\lambda_m^{ij}[k] = \frac{\delta_m^{ij} \cdot [\eta_i[k] \bar{\gamma}_{ij}[k] + (1 - \eta_i[k]) \gamma_{ij}^{\mu^-}]}{\sum_{j \in \Gamma^{-1}(i)} \delta_m^{ij} \cdot [\eta_i[k] \bar{\gamma}_{ij}[k] + (1 - \eta_i[k]) \gamma_{ij}^{\mu^-}]} \quad (13)$$

Where, $\min \{ q_m^i[k] + x_m^i[k], Q_m^i \cdot g_n^p[k] \}$ depicts the flows potentially departing from lane group m at time step k ; $\lambda_m^{ij}[k]$ is the percentage of traffic in lane group m going from link i to j . Therefore,

$\min \{ q_m^i[k] + x_m^i[k], Q_m^i \cdot g_n^p[k] \} \cdot \lambda_m^{ij}[k]$ reflects the flows potentially departing from link i to j in lane group m , and the summation of it over all lane groups in link i comes to Eq. (12). Assuming that a total of one unit flow is to depart from link i at time step k , $[\eta_i[k] \bar{\gamma}_{ij}[k] + (1 - \eta_i[k]) \cdot \gamma_{ij}^{\mu^-}]$ will be the amount of flows within that one unit to go to link j , and $\delta_m^{ij} \cdot [\eta_i[k] \bar{\gamma}_{ij}[k] + (1 - \eta_i[k]) \cdot \gamma_{ij}^{\mu^-}]$ will be the amount of flows going to link j by lane group m , and $\sum_{j \in \Gamma^{-1}(i)} \delta_m^{ij} \cdot [\eta_i[k] \bar{\gamma}_{ij}[k] + (1 - \eta_i[k]) \cdot \gamma_{ij}^{\mu^-}]$ will be the total amount of flows

departing from lane group m . Hence, $\lambda_m^{ij}[k]$ can be approximated with Eq. (13). Similarly, the percentage of normal arterial traffic in lane group m going from link i to j , $\bar{\lambda}_m^{ij}[k]$, can be approximated by Eq. (14):

$$\bar{\lambda}_m^{ij}[k] = \frac{\delta_m^{ij} \cdot \eta_i[k] \cdot \bar{\gamma}_{ij}[k]}{\sum_{j \in \Gamma^{-1}(i)} \delta_m^{ij} \cdot [\eta_i[k] \bar{\gamma}_{ij}[k] + (1 - \eta_i[k]) \gamma_{ij}^{\mu^-}]} \quad (14)$$

Then, the percentage of detour traffic in lane group m going from link i to j can be obtained with $(\lambda_m^{ij}[k] - \bar{\lambda}_m^{ij}[k])$. The estimation of $\lambda_m^{ij}[k]$, $\bar{\lambda}_m^{ij}[k]$, and $(\lambda_m^{ij}[k] - \bar{\lambda}_m^{ij}[k])$ will then be used to estimate the actual departing flows by hereafter equations.

In addition to Eq. (12), the actual number of vehicles departing from link i to link j at time step k is also constrained by the available storage space of the destination link j . Since the total flow to one destination link j may consist of several flows from different upstream links, this study assumes that the free storage space of link j allocated to accommodate upstream departing flow from link i is proportional to link i 's potential departing flow. Therefore, the actual departing flows from link i to link j at time step k is given by the following equation:

$$Q_{ij}[k] = \min \left\{ Q_{ij}^{pot}[k], \frac{Q_{ij}^{pot}[k]}{\sum_{i \in \Gamma(j)} Q_{ij}^{pot}[k]} \cdot s_j[k] \right\} \quad (15)$$

Where, $s_j[k]$ is the available space in link j at time step k , and $\frac{Q_{ij}^{pot}[k]}{\sum_{i \in \Gamma(j)} Q_{ij}^{pot}[k]}$ is the proportion of the available space in link j allocated to accommodate flows from link i .

Then, the flows actually departing from lane group m can be easily obtained by:

$$Q_m^i[k] = \sum_{j \in \Gamma^{-1}(i)} Q_{ij}[k] \cdot \delta_m^{ij} \quad (16)$$

Finally, the actual departing flow of normal arterial traffic from link i to link j at time step k is given by:

$$\bar{Q}_{ij}[k] = \sum_{m \in S_i^M} Q_m^i[k] \cdot \bar{\lambda}_m^{ij}[k] \quad (17)$$

And, the actual departing flow of detour traffic from link i to link j heading to on-ramp μ^- at time step k is given by:

$$Q_{ij}^\mu[k] = \sum_{m \in S_i^M} Q_m^i[k] \cdot (\lambda_m^{ij}[k] - \bar{\lambda}_m^{ij}[k]) \quad (18)$$

Flow conservation

The lane group based queues are advanced as follows:

$$x_m^i[k+1] = x_m^i[k] + q_m^i[k] - Q_m^i[k] \quad (19)$$

Queues outside the approach lanes due to overflows or blockages are advanced as follows:

$$\tilde{x}_m^i[k+1] = \tilde{x}_m^i[k] - q_m^i[k] + \sum_{j \in \Gamma^{-1}(i)} q_i^{arr}[k] \cdot [\eta_i[k] \bar{\gamma}_{ij}[k] + (1 - \eta_i[k]) \gamma_{ij}^{\mu^-}[k]] \cdot \delta_m^{ij} \quad (20)$$

Then, the total number vehicles queued at link i is computed as:

$$x_i[k+1] = \sum_{m \in S_i^M} (x_m^i[k+1] + \tilde{x}_m^i[k+1]) \quad (21)$$

The evolution of the total number of normal arterial vehicles at link i can be stated as:

$$\bar{N}_i[k+1] = \bar{N}_i[k] + \sum_{j \in \Gamma(i)} \bar{Q}_{ji}[k] - \sum_{j \in \Gamma^{-1}(i)} \bar{Q}_{ij}[k] \quad (22)$$

The evolution of the total number of detour vehicles present at link i can be stated as:

$$N_i^{\mu^-}[k+1] = N_i^{\mu^-}[k] + \sum_{j \in \Gamma(i)} Q_{ji}^{\mu^-}[k] - \sum_{j \in \Gamma^{-1}(i)} Q_{ij}^{\mu^-}[k] \quad (23)$$

And, the fraction of normal arterial traffic at link i is updated as:

$$\eta_i[k+1] = \frac{\bar{N}_i[k+1]}{\bar{N}_i[k+1] + N_i^{\mu^-}[k+1]} \quad (24)$$

Finally, one can compute the available storage space of link i as follows:

$$s_i[k+1] = N_i - \bar{N}_i[k+1] - N_i^{\mu^-}[k+1] \quad (25)$$

B. Freeway and Ramps

The macroscopic traffic flow model proposed by Messmer and Papageorgiou [34] was employed in this study to model traffic evolution for the freeway section. The key concept is to divide the freeway link into homogeneous segments, and update the flow, density, and speed within each segment at every time interval (ΔT). The model formulations can be found in [34] and will not be detailed in this paper.

As on-ramps and off-ramps function to exchange diversion flows between the freeway and arterial systems, this study has employed the lane-group-based concept to model their interactions. As illustrated in Fig. 4, one can model the on-ramp μ^- as a simplified arterial link with only one lane group and one downstream link. The only difference between an on-ramp and an arterial link is the departing process. Since the update step for freeway (ΔT) is usually larger than the one for arterial (Δt), this study has employed the approach by Van den Berg [30] to keep

consistency between the indices of time steps for the two systems (t is the time index for freeway and k is for arterial, and $k = l \cdot t$, $l = \Delta T / \Delta t$). Therefore, the actual flow that departs from on-ramp μ^- into freeway at time step k between $l \cdot t$ and $l \cdot (t+1) - 1$ is given by:

$$Q_{\mu^-}[k] = \frac{Q_{\mu^-}[t] \cdot \Delta T}{l}, \quad \forall k \in [l \cdot t, l \cdot (t+1) - 1] \quad (26)$$

and

$$Q_{\mu^-}[t] = \min \left(\frac{x^{\mu^-}[l \cdot t] + \sum_{k=lt}^{l(t+1)-1} q_{\mu^-}^{arr}[k]}{\Delta T}, Q_{\mu^-} \cdot R_{\mu^-}, Q_{\mu^-} \cdot \min[1, \frac{\rho^{jam} - \rho_{i+1,0}[t]}{\rho^{jam} - \rho_i^{crit}}] \right) \quad (27)$$

Where, $x^{\mu^-}[l \cdot t] + \sum_{k=lt}^{l(t+1)-1} q_{\mu^-}^{arr}[k]$ is the potential number of vehicles to merge into freeway mainline from on-

ramp μ^- at the freeway update time step t ; R_{μ^-} is the metering rate at on-ramp μ^- ; Q_{μ^-} is the discharge capacity of on-ramp μ^- ; ρ^{jam} is the jam density for freeway, $\rho_{i+1,0}[t]$ is the density of freeway segment immediate downstream of the on-ramp μ^- , and ρ_i^{crit} is the critical density of freeway link i where the incident occurs.

Similarly, the off-ramp could also be modeled as an arterial link if the upstream arrival process is modified properly, as shown in Fig. 4. The actual flow that enters off-ramp ν^+ at each arterial time step k between $l \cdot t$ and $l \cdot (t+1) - 1$ is given by:

$$q_{\nu^+}^{in}[k] = \frac{q_{\nu^+}^{in}[t] \cdot \Delta T}{l}, \quad \forall k \in [l \cdot t, l \cdot (t+1) - 1] \quad (28)$$

$$q_{\nu^+}^{in}[t] = \min \left\{ \rho_{i-1,N(i-1)}[t] \cdot v_{i-1,N(i-1)}[t] \cdot n_{i-1,N(i-1)} \cdot (\gamma_{\nu^+}^T + \beta_{\nu^+}^T \cdot Z_{\nu^+}^T), \frac{s_{\nu^+}[l \cdot t] + \sum_{k=lt}^{l(t+1)-1} \sum_{j \in \Gamma^{-1}(\nu^+)} Q_{\nu^+ j}[k]}{\Delta T} \right\} \quad (29)$$

Where, $\rho_{i-1,N(i-1)}[t]$, $v_{i-1,N(i-1)}[t]$, and $n_{i-1,N(i-1)}$ represent the density, speed, and number of lanes at the segment immediate upstream to the off-ramp ν^+ , respectively. $\gamma_{\nu^+}^T$ is the normal exit rate for off-ramp ν^+ during

control time interval T ; $Z_{\nu^+}^T$ is the diversion control rate to be determined during the control interval T ; $\beta_{\nu^+}^T$ is the driver compliance rate to the detour operation during control interval T ; Q_{ν^+} represents the discharge capacity of

off-ramp ν^+ , and $s_{\nu^+}[l \cdot t] + \sum_{k=lt}^{l(t+1)-1} \sum_{j \in \Gamma^{-1}(\nu^+)} Q_{\nu^+j}[k]$ is the available space at off-ramp ν^+ .

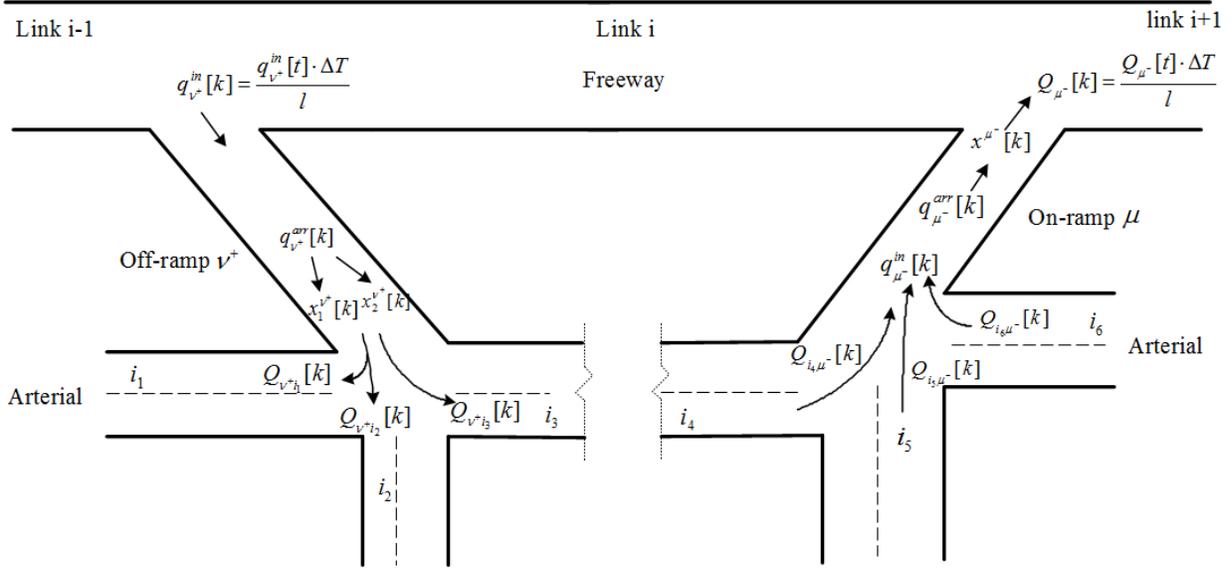


Fig. 4 Traffic flow exchange at on-off ramps

It should be mentioned that the integrated control process presented in this study has the flexibility to accommodate any new formulations for freeway or arterial flow dynamics [35-38].

IV. THE INTEGRATED CONTROL MODEL

Based on the above network flow formulations, this section presents a multi-objective control model to determine the set of control strategies that can efficiently explore the control effectiveness under different policy priorities between the target freeway and the available detour route.

A. Objective Function

Given the entire time horizon H for control, the first objective of the control model is to maximize the utilization of the parallel arterial so as to relieve congestion on the freeway mainline. This objective can further be stated as maximizing the total throughput of the freeway corridor during the incident management period by using the parallel arterial as the detour route. Since the throughput equals the total number of vehicles entering the

freeway link downstream of the on-ramp μ^- plus the total number of vehicles entering the arterial outgoing links, it can be stated as:

$$\max \quad \sum_{t=1}^H q_{i+1,0}[t] \cdot \Delta T + \sum_{k=1}^H \sum_{i \in S^{OUT}} q_i^{in}[k] \quad (30)$$

Where, $q_{i+1,0}[t]$ is the flow rate entering the freeway link $(i+1)$ downstream of the on-ramp μ^- ; S^{OUT} is the set of outgoing links in the arterial network.

The second objective function is to reflect the perspective of detour travelers, focusing on minimizing their total time on the detour route to ensure their compliance to the routing guidance. This objective is given by:

$$\min \quad \sum_{k=1}^H \left[\sum_{i \in S^U} N_i^{\mu^-}[k] + N_{\nu^+}^{\mu^-}[k] + N_{\mu^-}^{\mu^-}[k] \right] \cdot \Delta t \quad (31)$$

Where, $N_i^{\mu^-}[k]$, $N_{\nu^+}^{\mu^-}[k]$, and $N_{\mu^-}^{\mu^-}[k]$ represent the number of detour vehicles present at link i , off-ramp ν^+ , and on-ramp μ^- within the control area at time step k , respectively.

B. Decision Variables

The control variables need to be solved in the optimization formulation include:

$\{C^T, T \in H\}$: Common cycle length of the target arterial for each control interval;

$\{\Delta_n^T, \forall n \in S_N, T \in H\}$: Offset of intersection n for each control interval;

$\{G_{np}^T, \forall n \in S_N, p \in P_n, T \in H\}$: Green time for phase p of intersection n for each control interval;

$\{R_{\mu^+}^T, R_{\mu^-}^T, T \in H\}$: Metering rate at on-ramps μ^+ and μ^- for each control interval; and

$\{Z_{\nu^+}^T, T \in H\}$: Diversion rate at off-ramp ν^+ for each control interval;

C. Constraints

Representing the traffic state evolution along different parts of the traffic corridor, network formulations proposed in Section III constitute the principal constraints for the integrated control model. Moreover, the following constraints are common restrictions for the control decision variables:

$$C^{\min} \leq C^T \leq C^{\max}, \forall T \in H \quad (32)$$

$$G_{np}^{\min} \leq G_{np}^T < C^T, \forall n \in S_N, p \in P_n, T \in H \quad (33)$$

$$\sum_{p \in P_n} G_{np}^T + \sum_{p \in P_n} I_{np} = C^T, \forall n \in S_N, p \in P_n, T \in H \quad (34)$$

$$0 \leq \Delta_n^T < C^T, \forall n \in S_N, T \in H \quad (35)$$

$$R^{\min} \leq R_{\mu^+}^T, R_{\mu^-}^T \leq R^{\max}, T \in H \quad (36)$$

$$\beta_{v^+}^T \cdot Z_{v^+}^T + \gamma_{v^+}^T \leq Z^{\max}, T \in H \quad (37)$$

Where, C^{\min} and C^{\max} are the minimum and maximum for the cycle length, respectively; P_n is the set of signal phases at intersection n ; G_{np}^{\min} is the minimal green time for phase p of intersection n ; and I_{np} represents the clearance time for phase p of intersection n . R^{\min} and R^{\max} are the minimum and maximum metering rates at on-ramps, and Z^{\max} is the maximum percentage of traffic (including both detour and normal exiting) that can diverge from freeway to arterial.

Eq. (32) restricts the common cycle length to be between the minimal and maximal values. Eq. (33) requires that the green time for each phase should at least satisfy the minimal green time, and not exceed the cycle length. The sum of green times and clearance times for all phases at intersection n should be equal to the cycle length (see Eq. (34)). Furthermore, the offset of intersection n shall be constrained by Eq. (35), and lie between 0 and the cycle length. Eq. (36) limits the metering rates for on-ramps, and the diversion rate is bounded by Eq. (37).

Note that, the arterial traffic flow equations are not explicitly related to the signal control variables C^T , Δ_n^T , and G_{np}^T . To represent the signal status of phase p at each time step k , the binary variable $g_n^p[k]$ is employed to indicate whether or not the corresponding phase p is green. For a signal controller with a set of phases P_n shown in Fig. 5, we can employ the following equations to model relations between the phase status at time step k and signal control parameters:

$$(\delta_n^{p}[k] - 0.5) \cdot \text{mod}(k - \Delta_n^T, C^T) \leq (\delta_n^{p}[k] - 0.5) \cdot \sum_{j=1}^{p-1} (G_{nj}^T + I_{nj}) \quad (38)$$

$$p \in P_n, n \in S_N, T \in H, k \in T$$

$$(\delta_n^{\prime p}[k] - 0.5) \cdot \text{mod}(k - \Delta_n^T, C^T) > (\delta_n^{\prime p}[k] - 0.5) \cdot \left(\sum_{j=1}^{p-1} (G_{nj}^T + I_{nj}) + G_{np}^T \right) \quad (39)$$

$$p \in P_n, n \in S_N, T \in H, k \in T$$

$$g_n^p[k] = 1 - \delta_n^{\prime p}[k] - \delta_n^{\prime\prime p}[k], \quad p \in P_n, n \in S_N, k \in T \quad (40)$$

$\{\delta_n^{\prime p}[k], \delta_n^{\prime\prime p}[k], p \in P_n, n \in S_N, k \in T\}$ are a set of auxiliary 0-1 variables.

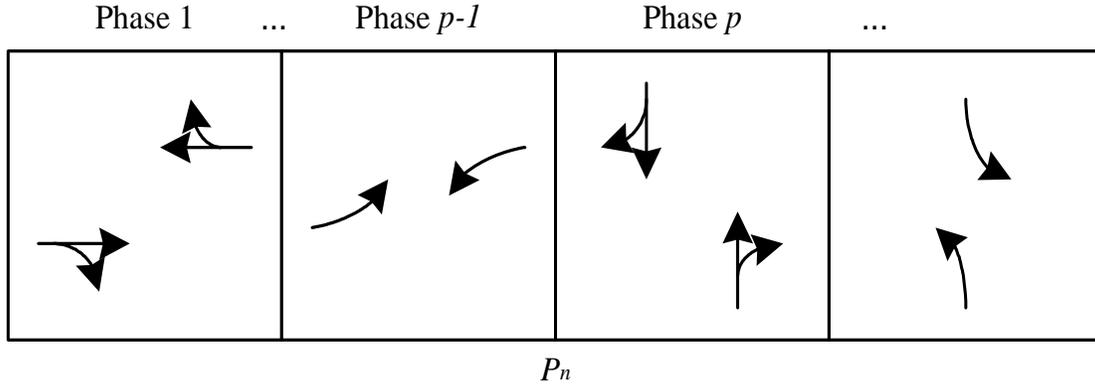


Fig. 5 A signal controller with a set of phases P_n

Other constraints include nonnegative constraints and initial values of the link state variables in the corridor network, which can be obtained from the on-line surveillance system to reflect the actual network condition preceding the onset of an incident. In summary, the mathematical description of the integrated corridor control problem is recapitulated as follows:

$$\min \quad \Phi(s) = \begin{bmatrix} f_1(s) \\ f_2(s) \end{bmatrix}$$

$$f_1 : - \left[\sum_{t=1}^H q_{i+1,0}[t] \cdot \Delta T + \sum_{k=1}^H \sum_{i \in S^{OUT}} q_i^{in}[k] \right]$$

$$f_2 : \sum_{k=1}^H \left[\sum_{i \in S^U} N_i^{\mu^-}[k] + N_{v^+}^{\mu^-}[k] + N_{\mu^-}^{\mu^-}[k] \right] \cdot \Delta t$$

$$s.t. \quad s : [C^T, \Delta_n^T, G_{np}^T, R_{\mu^+}^T, R_{\mu^-}^T, Z_{v^+}^T] \in S \quad (41)$$

Where, S denotes the feasible set defined by the network flow constraints and operational constraints.

V. SOLUTION APPROACH

Note that the formulations for the proposed integrated control model feature a bi-objective optimization framework. To apply the proposed models in practice, we have extended a GA-based heuristic by Gen and Cheng [39] and developed approximate solutions for each control interval during the entire optimization period. The proposed heuristic can identify the solution closest to the “ideally best point” of the multi-objective problem rather than to evaluate the entire *Pareto* solution set, which is quite time-consuming and poses a considerable cognitive burden on the traffic management decision makers.

In addition, considering the significant increase in the number of decision variables over different control intervals, application of the proposed large-scale, non-linear, and multi-objective control model remains challenging. Also, solving such a large-scale control system requires reliable projection of traffic conditions over the entire control horizon, which is also quite difficult due to the expected fluctuation of traffic flows and discrepancy of driver responses to control actions under non-recurrent congestion. To contend with the above critical issues, this study employs a rolling-horizon method for the GA-based heuristic, in which the model input and control strategy could be regularly updated to improve the computing efficiency and effectiveness under time-varying traffic conditions and potential system disturbance.

In this section, we will first detail the key components in the GA-based heuristic, and then illustrate the rolling horizon framework.

A. The GA-based Heuristic

The GA-based approach includes the following key components:

Objective function normalization

Note that in the proposed model, the first objective gives in number of vehicles, whereas the second objective is measured by vehicle-min scale. These two objectives cannot be compared or assigned weights directly. Hence, one needs to normalize the objective functions into a common satisfaction scale as follows:

$$\bar{f}_m(s) = \frac{f_m(s) - f_m^{\min}}{f_m^{\max} - f_m^{\min}} \in [0,1], s \in P, m = 1,2 \quad (42)$$

Where, $\bar{f}_m(s)$, f_m^{\min} , and f_m^{\max} are the normalized, minimum, and maximum value of objective function m .

Regret value computation

At each population in the evolution process of GA, the algorithm evaluates the performance of an individual solution by defining a regret value r , stated as follows:

$$r(s) = \left(\sum_{m=1}^M w_m \cdot \left| \bar{f}_m(s) - \bar{f}_m^* \right|^M \right)^{1/M}, s \in P \quad (43)$$

Where, $s \in P$ represents the solution s in the current population P ; w_m ($m = 1 \cdots M$) is the weight assigned to objective function m to emphasize its degree of importance; $\bar{f}_m(s)$ denotes the value of normalized objective function m corresponding to solution s , and \bar{f}_m^* is the value of normalized objective function m at the ideally best point, which will be zero according to Eq. (42). Considering the bi-objective model proposed in this study, Eq. (43) can be specified as:

$$r(s) = \sqrt{w_1 \cdot \left(\frac{f_1(s) - f_1^{\min}}{f_1^{\max} - f_1^{\min}} \right)^2 + w_2 \cdot \left(\frac{f_2(s) - f_2^{\min}}{f_2^{\max} - f_2^{\min}} \right)^2}, s \in P \quad (44)$$

Note that in Eq. (44), the smaller the regret value, the better the individual in the population.

Proxy ideally best point

Note that f_m^{\min} and f_m^{\max} in Eq. (44) for the proposed problem is difficult to obtain, so is the regret value.

Thus, we adopt a concept of proxy ideally best point to replace the real one. The proxy ideally best point is the best point corresponding to the current generation but not to the given problem, so it is easily obtained at each generation.

In this study, we use $f_m^{\min}(P) : \min\{f_m(s) \mid s \in P\}$ as the proxy ideally best point for objective function m

corresponding to the current population P , and f_m^{\max} was replaced by $f_m^{\max}(P) : \max\{f_m(s) \mid s \in P\}$. During the

evolution process, the proxy ideally best point will gradually approximate the real one. Thus, Eq. (44) is converted

into:

$$r(s) = \sqrt{w_1 \cdot \left(\frac{f_1(s) - f_1^{\min}(P)}{f_1^{\max}(P) - f_1^{\min}(P)} \right)^2 + w_2 \cdot \left(\frac{f_2(s) - f_2^{\min}(P)}{f_2^{\max}(P) - f_2^{\min}(P)} \right)^2}, s \in P \quad (45)$$

Fitness value computation

Finally, one needs to convert the regret value to the fitness value to ensure better individuals that have larger chance to evolve. For a minimization problem, the fitness value of an individual solution s in population P can be stated as:

$$eval(s) = \frac{r_{\max} - r(s) + \varepsilon}{r_{\max} - r_{\min} + \varepsilon}, s \in P \quad (46)$$

Where, r_{\max} and r_{\min} denote the maximum and minimum regret value in population P , respectively; ε is a positive value between 0 and 1 which functions to: 1) prevent Eq. (46) from zero division; and 2) adjust the selection behavior between fitness proportional selection and pure random selection [39].

Decoding for control variables

To generate feasible control parameters which satisfy the operational constraints, the following decoding scheme is employed:

- **Arterial signal control variables:** According to the phase structure shown in Fig. 5, within each control interval T , a total number of $NP_n + 1$ fractions ($\lambda_j^T, j = 1 \dots NP_n + 1$) are generated for the controller at intersection n from decomposed binary strings, where NP_n is the number of phases of intersection n . Those $NP_n + 1$ fractions are used to code the green times, cycle length, and offsets as shown by the following equations:

$$G_{np}^T = G_{np}^{\min} + (C^T - \sum_{j \in P_n} G_{nj}^{\min} - \sum_{j \in P_n} I_{nj}) \cdot \lambda_p^T \cdot \prod_{j=1}^p (1 - \lambda_{j-1}^T), p = 1 \dots NP_n - 1, n \in S_N \quad (47)$$

$$G_{np}^T = G_{np}^{\min} + (C^T - \sum_{j \in P_n} G_{nj}^{\min} - \sum_{j \in P_n} I_{nj}) \cdot \prod_{j=1}^p (1 - \lambda_{j-1}^T), p = NP_n, n \in S_N \quad (48)$$

$$C^T = C_{\min} + (C_{\max} - C_{\min}) \cdot \lambda_{NP}^T \quad (49)$$

$$\Delta_n^T = (C^T - 1) \cdot \lambda_{NP+1}^T \quad (50)$$

Eq. (49) constrains the random cycle lengths generated through the binary string within the maximum and minimum allowable cycle lengths. Using the cycle length generated above, Eq. (50) would result in an offset value

that lies between 0 and the cycle length minus one. The green times are assigned to each phase within a feasible range by Eqns. (47) and (48), in which λ_0 was set to zero to accommodate the case when $j = 1$.

- **Diversion and metering rates:** The following equations constrain the random diversion and metering rates generated within the maximum and minimum allowable range:

$$Z_{v^+}^T = \frac{(Z^{\max} - \gamma_{v^+}^T)}{\beta_{v^+}^T} \cdot \lambda_{v^+}^T \quad (51)$$

$$R_{\mu^+}^T = R^{\min} + (R^{\max} - R^{\min}) \cdot \lambda_{\mu^+}^T \quad (52)$$

$$R_{\mu^-}^T = R^{\min} + (R^{\max} - R^{\min}) \cdot \lambda_{\mu^-}^T \quad (53)$$

Where, $\lambda_{v^+}^T$, $\lambda_{\mu^+}^T$ and $\lambda_{\mu^-}^T$ are fractions generated through the decomposed binary string over each control interval.

B. The Rolling-horizon Approach

Rolling-horizon approach is a common practice for making decisions in a dynamic and stochastic environment. One key issue for employing a rolling-horizon framework in traffic control is to keep the consistency between the variation of arterial signal timings and the update of control time interval. The following two types of strategies are commonly employed in the literature: 1) arterial signal timings are represented with G/C (green time/cycle length) ratios and updated at every constant time interval, or 2) a constant network cycle length is pre-set to keep consistency with the control update interval, and green splits as well as offsets are optimized under the given cycle length. However, some limitations embedded in those approaches may limit their applications:

- The system implementation can not be based only on the G/C ratios for arterial intersections. It still needs an additional interface with a compatible microscopic local control controller to determine the resulting signal phasing, green times, and offsets;
- A pre-set network cycle length may not be able to accommodate the traffic fluctuation under incident conditions; and
- A constant control update time interval may not be sufficiently responsive to the variation of signal control parameters, thus may cause the loss of some phases.

To address the above issues, this study has employed the following rolling horizon framework (see Fig. 6):

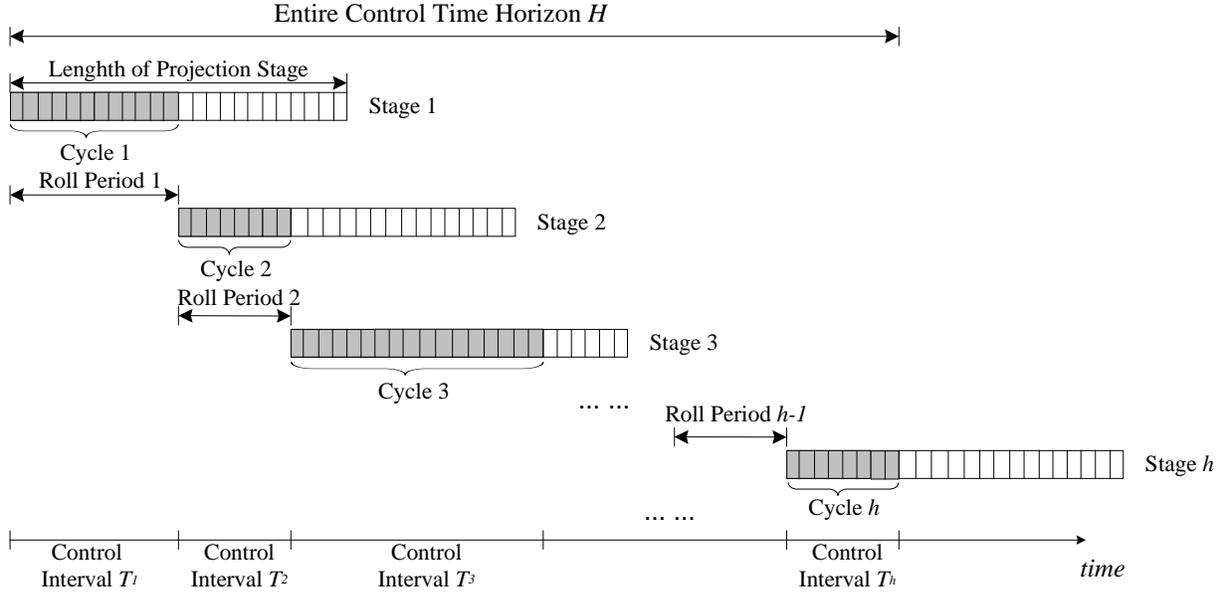


Fig. 6 Illustration of the rolling horizon framework

- Control policies are calculate over each projection stage, as shown in Fig. 6, but implemented only for the control interval T (head section of each stage), which is an integer multiple of the cycle length calculated in that stage; and
- Once the control plan is implemented, the state of the traffic within the corridor network is updated with real-time measurements from the surveillance system, and the optimization process starts all over again with the prediction and the control horizon shifted forward by one control interval.

Note that, in real-world applications, one needs to exercise care in selecting the lengths of the prediction horizon and the control interval based on the computational complexity and the controller accuracy.

VI. EXPERIMENTAL TESTS

A. Experimental Test Network

To evaluate the performance of the proposed model, this study employs a corridor segment along I-95 Northbound from Washington DC to Baltimore (shown in Fig. 7) for experimental analysis. Assuming that an incident occurs on the freeway mainline segment (between node 26 and 44), traffic will detour to MD198 and then follow MD216 back to the freeway. The proposed control model will update the control measures, including diversion rate at node 27, signal timings at intersections along MD198 and 216 (nodes 68, 69, 65, 67, and 99), and

metering rate at node 26 and 43. The entire test period is designed to cover 35 minutes, which consists of the following 3 periods: 5 min for normal operations (no incident), 20 min with incident, and 10 min recovery period (incident cleared). A total of four scenarios are designed as follows for the experiment (see Table 2 for volume levels):

- Scenario I: Volume level-I with 1 lane blocked due to incident;
- Scenario II: Volume level-I with 2 lanes blocked due to incident;
- Scenario III: Volume level-II with 1 lane blocked due to incident; and
- Scenario IV: Volume level-II with 2 lanes blocked due to incident.

B. Key Model Parameter Settings

Within the control area shown in Fig.7, I-95 mainline has 4 lanes in the northbound direction. On the detour routes, the off-ramp from I-95 North to MD198 East has 2 lanes, and MD198 East is an arterial street with 3 lanes in each direction. MD216 is an arterial street with 2 lanes in each direction, and the on-ramp from MD216 to I-95 North has 1 lane. The lane channelization at each intersection is shown in Table 3 and the phase diagram is summarized in Table 4.

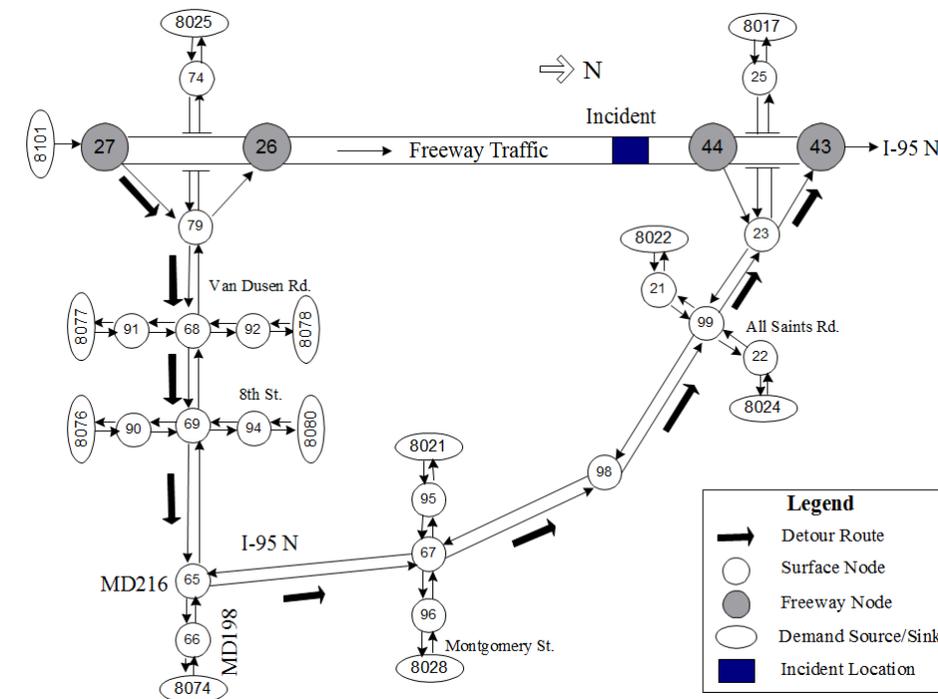


Fig. 7 Layout of the corridor segment for experimental test of the integrated control model

TABLE 2 Volume levels for the experimental test

Demand Entries	Level I (vph)	Level II (vph)
8101	4680	7800
8025	614	1024
8017	564	940
8077	554	924
8078	725	1208
8076	200	400
8080	210	384
8074	550	916
8021	200	400
8028	246	510
8022	187	312
8024	390	684

TABLE 3 Lane channelization at intersection approaches of the detour route

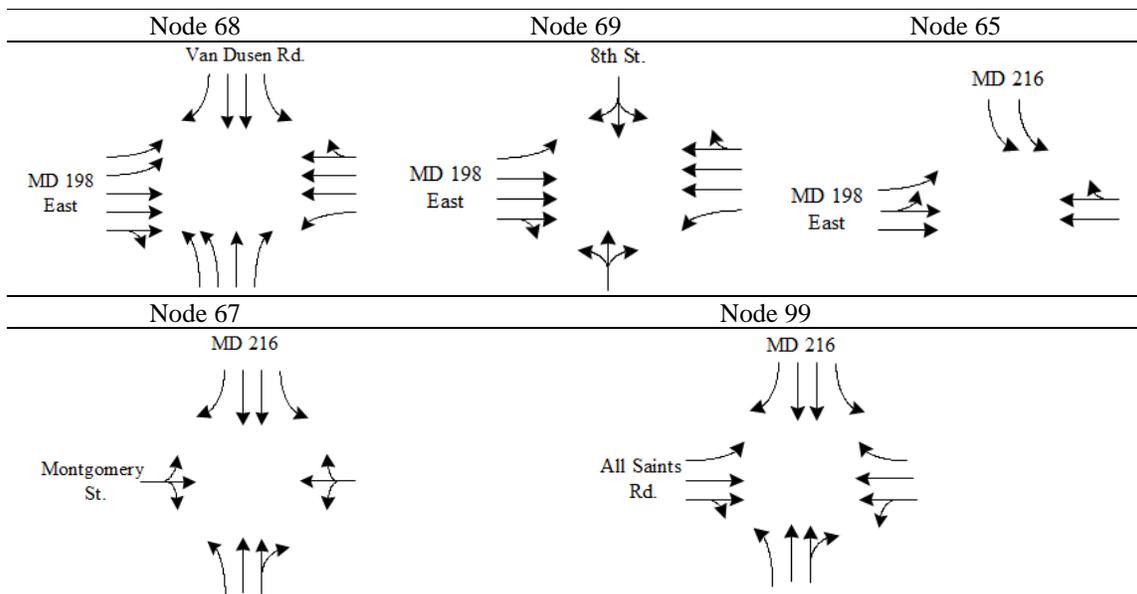
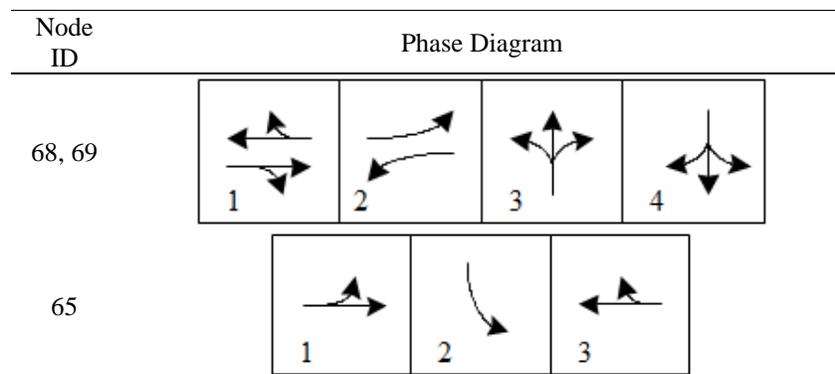
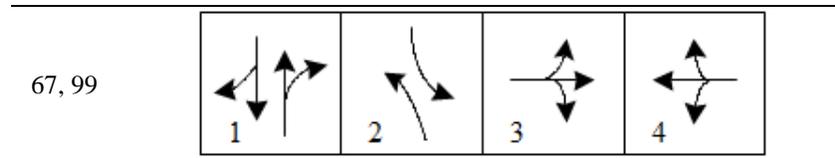


TABLE 4 Phase diagram of intersections along the detour route





All other parameters related to the network flow models, arterial signals, and solution algorithm in the test are summarized in Table 5:

TABLE 5 Key model parameters in the experimental test

Parameters*	Values
<u>Traffic flow model parameters</u>	
Δt , ΔT (in secs)	1, 5
Freeway segment length (in ft)	800
ρ^{jam} , ρ^{min} (in veh/mile/lane)	210, 20
Free flow speed at freeway, ramps, arterials (in mph)	65, 45, 50
Minimum speed corresponding to jam density (in mph)	5
Link discharge capacity for freeway, ramps, arterials (in vplph)	2200, 1900, 1800
Average vehicle length (in ft)	24
Freeway model parameters: α_f	1.78
Freeway model parameters: τ (in secs)	27
Freeway model parameters: η (in mile ² / h)	6
Freeway model parameters: κ (in veh/mile/lane)	21
Normal exiting rate at the off-ramp to MD198 East, $\gamma_{v^+}^T$	0.0875
Driver compliance rates to the detour operation, $\beta_{v^+}^T$ (if the detour travel time is less or comparable to the freeway travel time)	100%
<u>Arterial signal parameters</u>	
C^{min} , C^{max} at arterial intersections (in secs)	60, 160
G_{np}^{min} , I_{np} at arterials intersections (in secs)	7, 5
Minimum and maximum ramp metering rates: R^{min} , R^{max}	0.1, 1.0
Maximum diversion rate Z^{max}	0.25
<u>Parameters in solution algorithm</u>	
Weights of importance w_1 / w_2	Assigned from 10/0 to 0/10 at an increment of 1.
Population size in GA	50
Maximum number of generation in GA	200
Crossover probability in GA	0.5
Mutation probability in GA	0.03
Fitness selection parameter in GA: ε	0.1
Length of the projection stage in rolling horizon framework (in minutes)	4
Control update time interval T	one cycle length in each projection stage

C. Results

This study has implemented the proposed model and the solution procedure in Visual C++ 2005, with the embedded GA-based heuristic coded by the MIT GA C++ Library v.2.4.6 [40]. The control plans obtained from the proposed model will be evaluated through the following steps:

- Step I - Evaluate the performance of the proposed model with systematically varied weights to provide operational guidelines for decision makers in best weighting importance between both control objectives under each given scenario;
- Step II - With a set of properly selected weights from Step I, compare the model performance with the following two control strategies:
 - A. No control (just close the incident upstream on-ramp);
 - B. Diversion control with rates determined by a static user-equilibrium (UE) assignment between freeway and arterial, and re-timing of arterial signals with TRANSYT-7F based on volumes from the assignment results. On-ramp metering is operated with ALINEA.
- Step III - Evaluate the performance of the proposed control model with the diversion compliance rates varying at different levels.

The microscopic simulator CORSIM was employed as an unbiased evaluator for model performance. To overcome the stochastic nature of simulation results, an average of 30 simulation runs has been used.

Step-I: Weight Assignment

Fig. 8 summarizes the performance of the proposed model under different scenarios and weights of importance between control objectives. One can observe the following primary findings:

- For Scenario I, the performance of the bi-objective model is not sensitive to the weight variation, as shown in Fig. 8a. This is probably due to the fact that the existing capacity of the freeway can accommodate the demand in the Volume-I level without detour operations, and the ideally best point for both control objectives was reached. The slight fluctuation of the system objective function values is probably due to the convergence of GA within different control intervals;
- For Scenario II, the performance of the bi-objective model is not sensitive within a specific range. For example, the performance of the model seems quite stable as long as $w_1 > w_2$, as shown in Fig. 8b.

- That is probably due to the fact that the under-saturated arterial can accommodate sufficient detour traffic volume as long as the freeway system is given the priority. However, when $w_1 \leq w_2$, the total corridor throughput exhibits a dramatic drop (from 2808 vehs to 2680 vehs) due to the priority switching from the freeway to the arterial. When the arterial is given the highest priority (0/10), the corridor throughput will be at the lowest level (2512 vehs);
- For Scenarios III and IV, the performance of the model is sensitive to every step of the weight adjustment between objective functions (see Figs. 8c and 8d). Every improvement of the performance for one objective will be at the cost of the other.

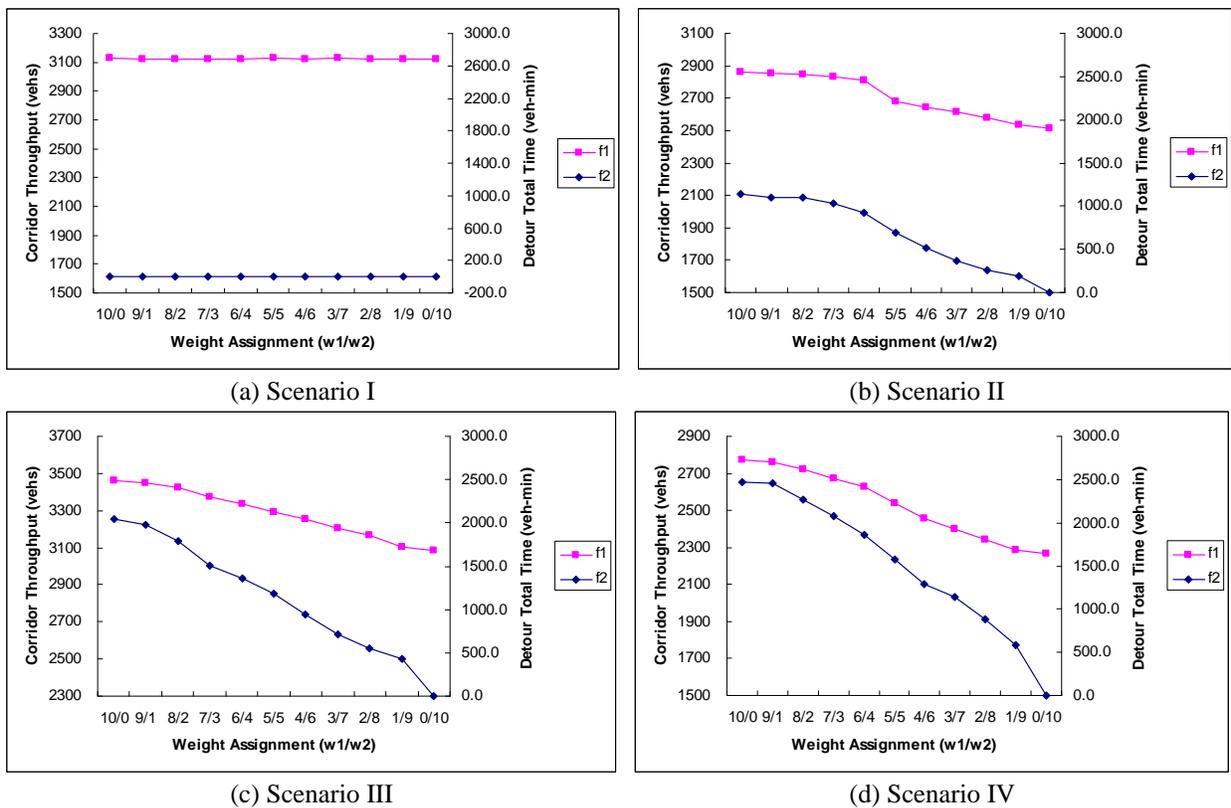


Fig. 8 System MOE changes under different weight assignment between control objectives

To further assist traffic operators in best weighting the importance between both control objectives, this section has also investigated the time-varying travel time patterns on both the detour route and the freeway mainline during the control period under different scenarios (see Fig. 9 – Fig. 11) except Scenario I (no detour in Scenario I). The following findings can be reached:

- With the weight assignment changing from $w_1 / w_2 = 10/0$ to $w_1 / w_2 = 0/10$, the ratio of detour travel time to freeway travel time decreases under all scenarios;
- The commonly used single control objective of maximizing the total corridor throughput (i.e. $w_1 / w_2 = 10/0$) may result in a significant unbalance of travel time between the detour route and the freeway mainline which could cause unacceptable driver compliance rates and degrade the control performance; and
- There exists a threshold in the weight assignment for each scenario, below which the assumed level of driver compliance rates can be achieved. For example, this case study assumes a 100% driver compliance rate if the detour travel time is less than or comparable to the freeway travel time, which indicates that the weight assignment must be set at a critical value to ensure the ratio of detour travel time to freeway travel time is less than or around 1.0. For Scenario I, one can set $w_1 / w_2 = 10/0$ to maximize the utilization of residual freeway capacity without detour operations. For Scenario II (see Fig. 9), one can still set $w_1 / w_2 = 10/0$ to fully utilize the available capacity in the arterial while keeping a high level of driver compliance rates. For Scenario III (see Fig. 10), one needs to set $w_1 / w_2 = 6/4$ or lower to ensure acceptable driver compliance rates. For Scenario IV (see Fig. 11), one needs to set $w_1 / w_2 = 5/5$ or lower to ensure acceptable driver compliance rates.

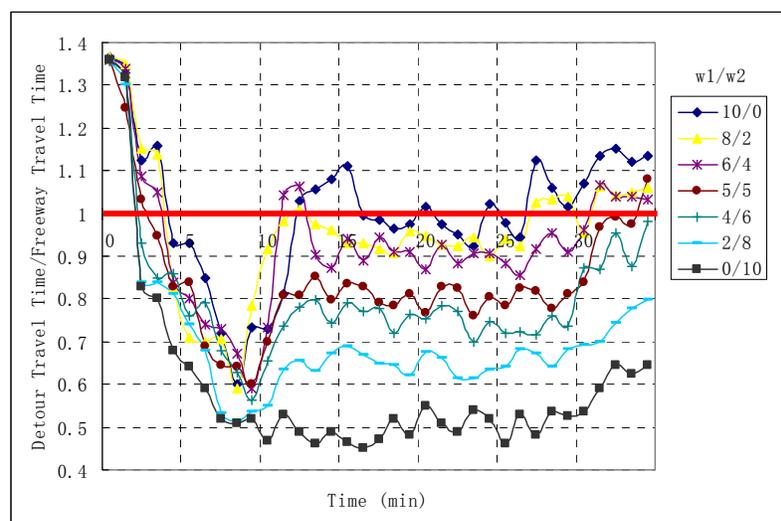


Fig. 9 Time-varying ratio of detour travel time to freeway travel time with different weight assignment (Scenario II)

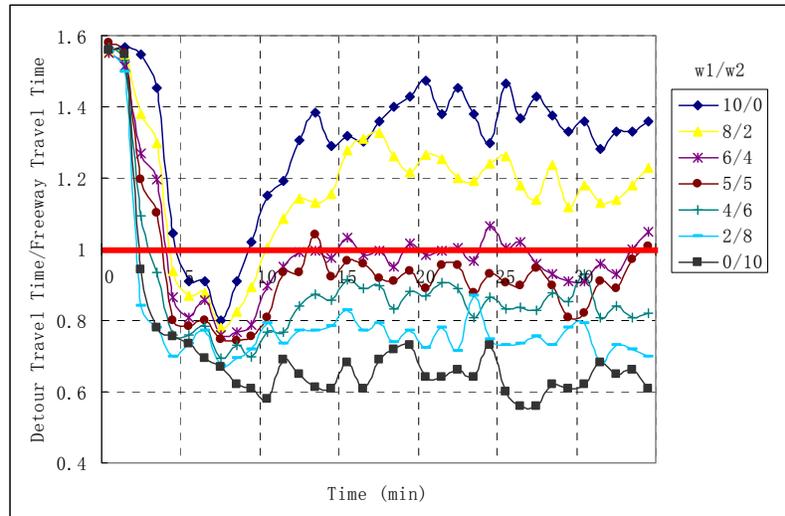


Fig. 10 Time-varying ratio of detour travel time to freeway travel time with different weight assignment (Scenario III)

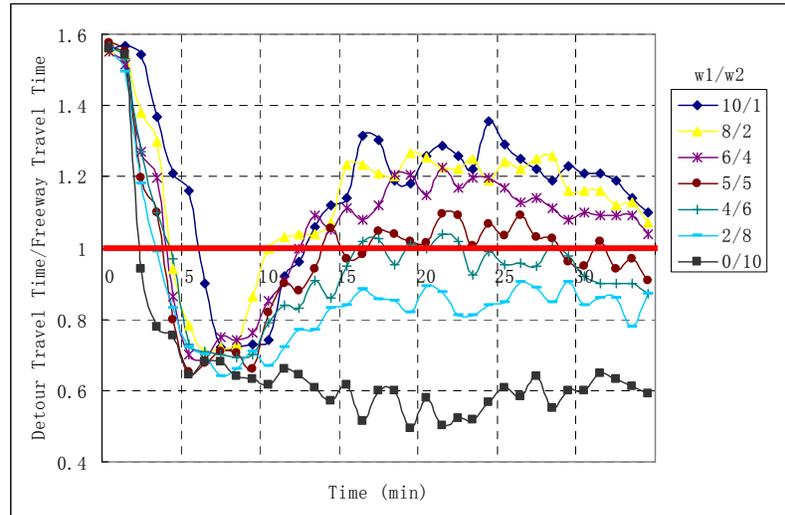


Fig. 11 Time-varying ratio of detour travel time to freeway travel time with different weight assignment (Scenario IV)

Step-II: Comparative Analysis of Model Performance

This study has also compared the performance of the proposed model with other incident management strategies with respect to the total corridor throughput increases and the total spent time savings under all scenarios.

The control strategies for comparison are:

- Control A: The base line - no control (just close the incident upstream on-ramp);
- Control B: Diversion control with rates determined by a static user-equilibrium (UE) assignment between the freeway and arterial, and re-timing of arterial signals with TRANSYT-7F based on volumes from the assignment results. On-ramp metering is operated with ALINEA.

Based on the analysis results from Step I, the weights of control objectives for the four test scenarios are set as follows to ensure acceptable compliance rates to detour operations:

- Scenario I: $w_1 / w_2 = 10/0$;
- Scenario II: $w_1 / w_2 = 10/0$;
- Scenario III: $w_1 / w_2 = 6/4$;
- Scenario IV: $w_1 / w_2 = 5/5$;

To consider the randomness caused by different initialization values of GA and its impact on the on-line control performance, for each incident scenario we have run the GA optimization process multiple times to yield the MOEs and take the mean value for comparison with other strategies. We have found the standard deviations for the accumulative throughput increases or total time savings are trivial, however in some control intervals GA won't converge to a reliable solution before the time window rolling to next stage. In this case, we take the control plan from the previous control interval. Since only less than 3% among all cases has this problem, it will not affect the applicability of the GA for on-line use.

Figs. 12-15 illustrate the comparison results. The following findings can be reached:

- The proposed model can outperform Control A and Control B for all scenarios in terms of both total time savings and total throughput increases at the assumed level of driver compliance rates.
- In Scenario I (see Fig. 12), since the existing capacity of the freeway can accommodate traffic without detour operations, the proposed model outperforms Control A probably due to the fact that the proposed model can produce slightly better signal timings in the arterial than TRANSYT-7F under light traffic conditions. Control B, however, has exhibited its performance worse than Control A, which is caused by the extra amount of traffic detoured to the arterial set by the static UE.
- In Scenario II (see Fig. 13), the proposed model compared with Control A, exhibits a substantial improvement since it aims to maximize the total corridor throughput ($w_1 / w_2 = 10/0$), which also results in a relatively low total spent time. However, the improvement over Control B is relatively low, that is probably due to the static UE employed in Control B which can also provide good utilization of the excessive capacity in the arterial. and

- In Scenarios III and IV (see Figs. 14 – 15), the proposed model significantly outperforms both Control A and Control B due to its integrated control function and the embedded traffic flow equations which are capable of capturing the evolution of detour traffic along the ramps and surface streets as well as the resulting local bottlenecks under saturated conditions.

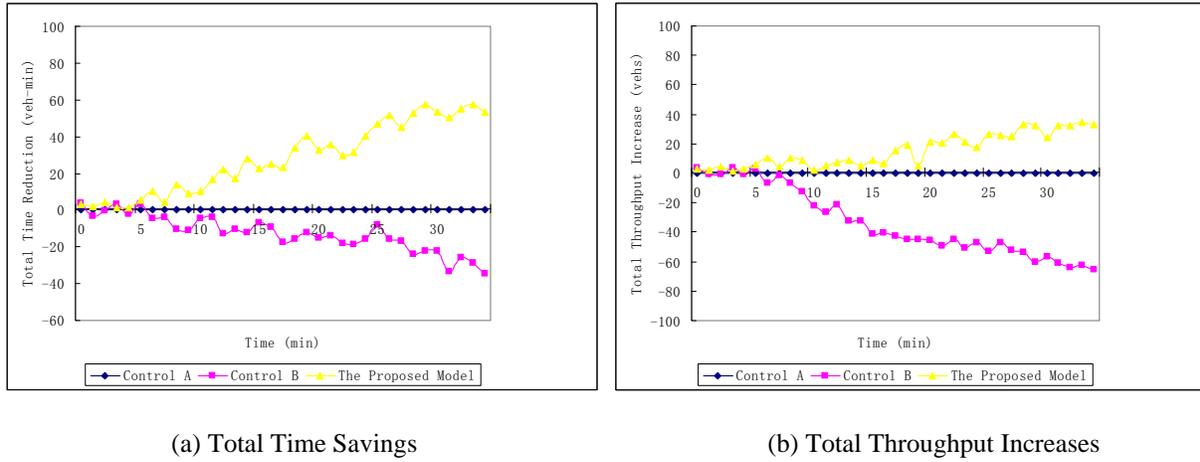


Fig. 12 Time-varying control performance comparison (Scenario I)

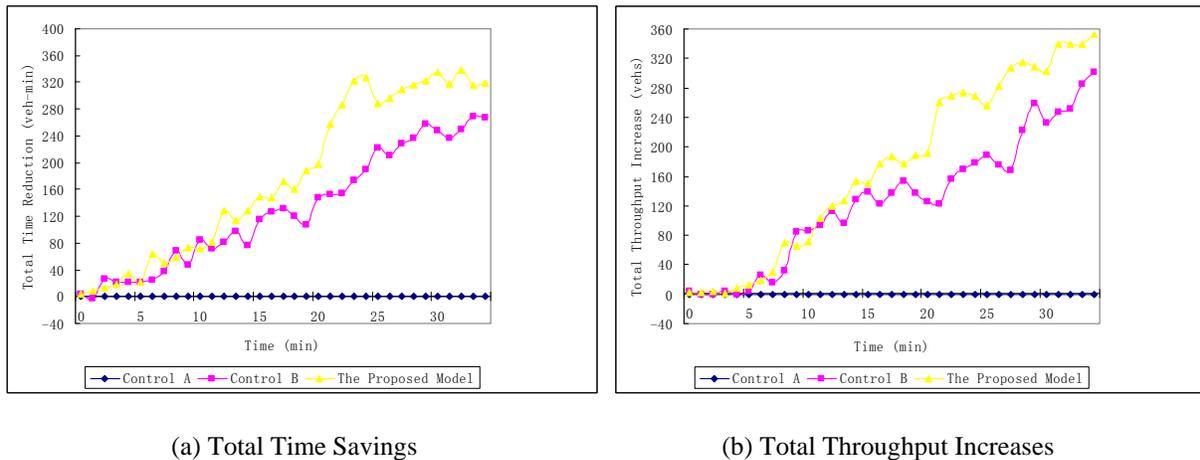
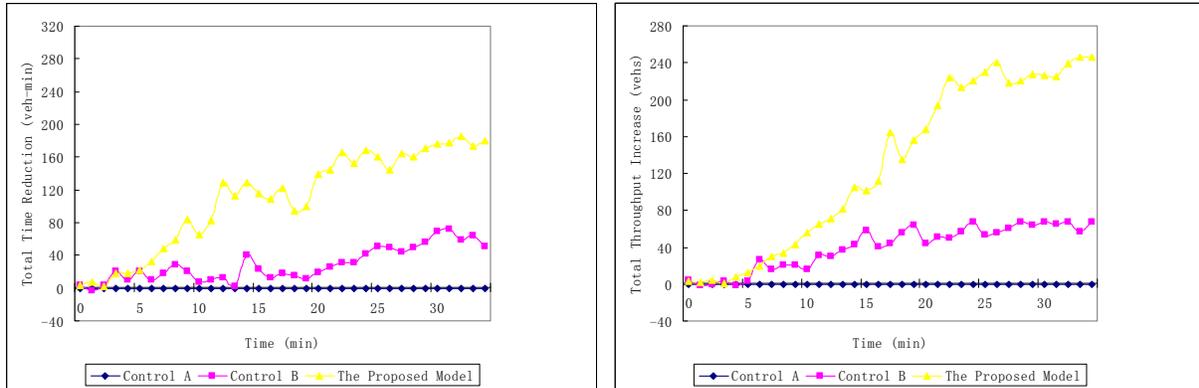


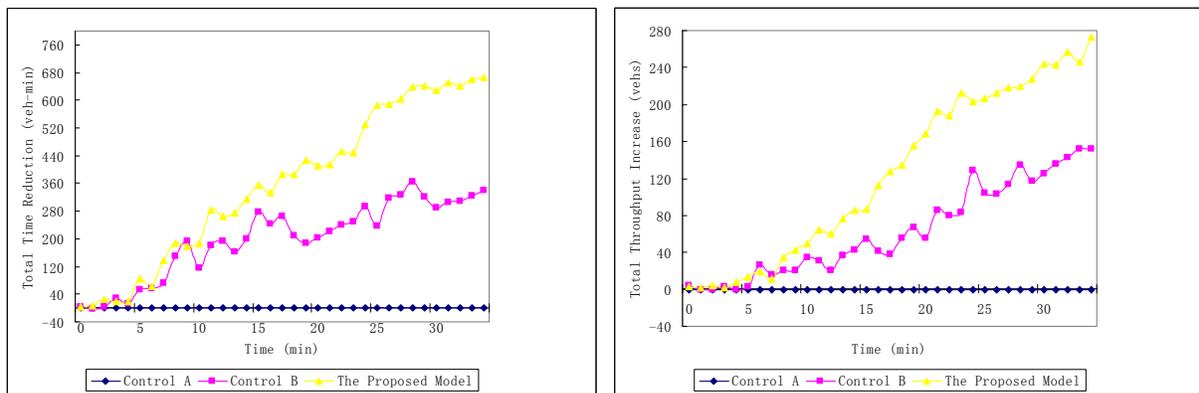
Fig. 13 Time-varying control performance comparison (Scenario II)



(a) Total Time Savings

(b) Total Throughput Increases

Fig. 14 Time-varying control performance comparison (Scenario III)



(a) Total Time Savings

(b) Total Throughput Increases

Fig. 15 Time-varying control performance comparison (Scenario IV)

Step-III: Sensitivity Analysis on Diversion Compliance

Conditioned on a 100% level of diversion compliance rates, the proposed integrated control model outperforms other control strategies with respect to both the total spent time savings and the total corridor throughput increases, as illustrated in Step-II analysis. However, during real-world operations, driver behavioral patterns are usually subject to time-varying fluctuations. Therefore, the sensitivity of the control performance with respect to the variation of diversion compliance rate needs to be investigated.

To address the above critical issue, this section has evaluated the performance of the proposed model under two previously designed experimental scenarios (Scenarios II and IV), with the diversion compliance rates varying at the 95%, 90%, 85%, 80%, and 70% levels, respectively. Table 6 has summarized the results of the sensitivity analyses.

TABLE 6 Sensitivity analysis of model performance with respect to the variation of diversion compliance rates

Level of Diversion Compliance	Incident Scenario	System MOEs			
		Total Time Spent Savings (veh-min)	Improvement over Control B	Total Throughput Increases (vehs)	Improvement over Control B
100%	Scenario II	319.14	19.0%	352	17.7%
	Scenario IV	667.03	95.6%	273	79.6%
95%	Scenario II	303.82	23.8%	337	24.3%
	Scenario IV	630.00	89.5%	254	77.0%
90%	Scenario II	275.04	21.4%	308	22.5%
	Scenario IV	500.99	104.9%	216	80.9%
85%	Scenario II	242.25	26.7%	289	28.0%
	Scenario IV	444.00	113.4%	198	81.4%
80%	Scenario II	199.21	30.6%	250	32.9%
	Scenario IV	264.00	120.4%	161	86.2%
70%	Scenario II	179.36	33.8%	232	41.0%
	Scenario IV	220.00	127.0%	139	90.7%

As illustrated in Table 6, the performance of the integrated control model declines with the decrease of the level of diversion compliance rates. For example under Scenario II, the savings of total spent time drop from 319.14 veh-min to 179.39 veh-min with the decrease of diversion compliance level from 100% to 70%, and the increases of total corridor throughput decline from 352 to 232 similarly. Such patterns can also be observed under the incident Scenario IV.

However, the relative percentage of improvement over Control B seems not to be sensitive to the decrease in the diversion compliance rates (see columns 4 and 6 in Table 6), which has indicated the potential for an application of the proposed model in the traffic environment with significant discrepancy in driver behavioral patterns.

VII. CONCLUSIONS

This study has presented an integrated model for design of effective control strategies for freeway corridor under non-recurrent congestion. To capture the critical operational issues at different components of the freeway corridor as well as to ensure computing efficiency, this study has proposed a set of macroscopic traffic flow models which can precisely model and predict the traffic evolution along the freeway mainline, arterial link, and on-off ramps, especially the local bottlenecks caused by the demand surge due to diversion. Aiming at maximizing the utilization of the available corridor capacity via detour operations, but not incurring excessive congestion for detour traffic on the arterials and ramps, this study proposed a multi-objective framework that can efficiently explore the control effectiveness under different priority policies between the target freeway and available detour routes for

different demand patterns and incident severity levels. A GA-based heuristic integrated with a rolling horizon framework has been employed to yield reliable solutions for the proposed model in real-time applications.

Experimental tests with a corridor segment along I-95 Northbound under various traffic conditions and incident levels have indicated that:

- For under-saturated scenarios, the performance of the bi-objective model is not quite sensitive to the weight assignment. Increasing the total corridor throughput by detouring traffic will not excessively degrade the arterial traffic conditions. Therefore, maximizing the total corridor throughput in those scenarios will be an effective decision;
- For oversaturated scenarios, the entire system performance becomes sensitive to the weight assignment. A single control objective of maximizing the total corridor throughput tends to favor the freeway mainline by detouring traffic to the arterial, which may cause over-congestion in the arterial and discourage the drivers to follow the detour guidance. In such scenarios, decision makers need to carefully set the weight assignment, based on the corridor network structure and driver behavioral characteristics so as to maximize the utilization of the corridor capacity while maintaining the acceptable diversion compliance rate;
- With properly selected weights of importance between control objectives, the proposed model outperforms the state-of-practice incident management strategies with better use of available corridor capacity and less total travel times, especially under saturated conditions; and
- The advantages of the proposed model over other control strategies are insensitive to the driver behavioral patterns and their responses to the control strategies.

Note that this paper has presented the evaluation results with a limited stretch of freeway corridor. More extensive experiments or field tests will be conducted to assess the effectiveness of the proposed model for decision support under various demand patterns, incident levels, turning proportions, and geometry configurations prior to its extension to control multiple segments of a corridor network.

REFERENCE

- [1] Wattleworth, J.A., Peak-period control of a freeway system—some theoretical considerations. Ph.D. Dissertation, Northwestern University, 1963.
- [2] Yuan, L. S., J. B. Kreer, “Adjustment of freeway ramp metering rates to balance entrance ramp queues,” *Transportation Research*, vol. 5, pp.127–133, 1971.
- [3] Yagar, S., “Metering freeway access,” *Transportation Quarterly*, 43, 215-224, 1989.
- [4] Papageorgiou, M., J.-M. Blosseville, H. Hadj-Salem, “Modeling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris - Part II: Coordinated on-ramp metering,” *Transportation Research* 24A, 5, 361-370, 1990.
- [5] Stephanedes, Y., and K. K. Chang, “Optimal ramp metering control for freeway corridors,” in *Proceedings of the 2nd International Conference on Applications of Advanced Technology in Transportation*, Minneapolis, Minnesota, 1991.
- [6] Papageorgiou, M., H. Hadj-Salem, and J.-M. Blosseville, “ALINEA: A local feedback control law for on-ramp metering,” *Transportation Research Record* 1320, 58-64, 1991.
- [7] Zhang, H., S. Ritchie, W. Recker, “Some general results on the optimal ramp metering control problem,” *Transportation Research* 4C, 51-69, 1996.
- [8] Zhang, H.M., W. Recker, “On optimal freeway ramp control policies for congested traffic corridors,” *Transportation Research* 33B, 417-436, 1999.
- [9] Kwon, E., Nanduri, S., Lau, R., Aswegan, J., “Comparative analysis of operational algorithms for coordinated ramp metering,” *Transportation Research Record*, n 1748, p 144-152, 2001.
- [10] Chang, T.H., Z.Y. Li, “Optimization of mainline traffic via an adaptive coordinated ramp metering control model with dynamic OD estimation,” *Transportation Research* 10C, 99-120, 2002.
- [11] Zhang, L., D.M. Levinson, “Optimal freeway ramp control without origin–destination information,” *Transportation Research* 38B, 869-887, 2004.
- [12] Perrine, K.A., Y. Lao, Y. Wang, “Area-wide System for Coordinated Ramp Meter Control,” *TRB 88th Annual Meeting Compendium of Papers*, paper 09-2243, 2009.
- [13] Payne, H. J., “Models of freeway traffic and control,” *Simulation Conference Proceedings* 1, 51-61, 1971.

- [14] Papageorgiou, M., "A hierarchical control system for freeway traffic," *Transportation Research* 17B, 251-261, 1983.
- [15] Stephanedes, Y., and K. K. Chang, "Optimal control of freeway corridors," *ASCE Journal of Transportation Engineering* 119, 504-514, 1993.
- [16] Messmer, A., M. Papageorgiou, "Automatic control methods applied to freeway network traffic," *Automatica*, 30, 691-702, 1994.
- [17] Kotsialos, A., M. Papageorgiou, M. Mangeas, H. Haj-salem, "Coordinated and integrated control of motorway networks via non-linear optimal control," *Transportation Research* 10C, 65-84, 2002.
- [18] Papageorgiou, M., "Dynamic modeling, assignment, and route guidance in traffic networks," *Transportation Research* 24B, 471-495, 1990.
- [19] Messmer, A., M. Papageorgiou, "Route diversion control in motorway networks via nonlinear optimization," *IEEE Transaction on Control System Technology*, 3, 144-154, 1995.
- [20] Iftar, A., "A decentralized routing controller for congested highways," In *Proceedings of 34th Conference on Decision and Control*. New Orleans, LA, USA, pp. 4089-4094, 1995.
- [21] Hawas, Y., H.S. Mahmassani, "A decentralized scheme for real-time route guidance in vehicular traffic networks," In *Proceedings of 2nd World Congress on Intelligent Transport Systems*, pp. 1956-1963, 1995.
- [22] Wu, J., and G.L. Chang, "Heuristic method for optimal diversion control in freeway corridors," *Transportation Research Record* 1667, 8-15, 1999.
- [23] Ben-Akiva, M., Bottom, J., Ramming, M.S., "Route guidance and information systems," *Journal of Systems and Control Engineering*, v 215, n 4, p 317-324, 2001.
- [24] Paz, A., Peeta, S., "Behavior-consistent real-time traffic routing under information provision," *Transportation Research Part C*, v 17, n 6, p 642-661, 2009.
- [25] Cremer, M., S. Schoof, "On control strategies for urban traffic corridors," In *Proceedings of IFAC Control, Computers, Communications in Transportation*, Paris, 1989.
- [26] Zhang, Y., A. Hobeika, "Diversion and signal re-timing for a corridor under incident conditions," presented at *77th Annual Meeting of Transportation Research Board*, Washington, DC, 1997.
- [27] Chang, G. L., P. K. Ho, C. H. Wei, "A dynamic system-optimum control model for commuting traffic corridors," *Transportation Research* 1C, 3-22, 1993.

- [28] Papageorgiou, M., "An integrated control approach for traffic corridors," *Transportation Research* 3C, 19-30, 1995.
- [29] Wu, J., and G.L. Chang, "An integrated optimal control and algorithm for commuting corridors," *International Transactions on Operations Research* 6, 39-55, 1999.
- [30] Van den Berg, De Schutter, A. Hegyi, B. and J. Hellendoorn, "Model predictive control for mixed urban and freeway networks," *Transportation Research Record*, 1748, 55-65, 2001.
- [31] Liu, Y., J., Yu, G.L., Chang, S., Rahwanji, A lane-group based macroscopic model for signalized intersections account for shared lanes and blockages. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 639-644, 2008.
- [32] Liu, Y., Chang, G.-L. An arterial signal optimization model for intersections experiencing queue spillback and lane blockage. *Transportation Research Part C* (2010), doi:10.1016/j.trc.2010.04.005.
- [33] Ben-Akiva, M., Development of a deployable real-time dynamic traffic assignment system, Task D Interim report: analytical developments for DTA system, MIT, Cambridge, MA, 1996.
- [34] Messmer, A., Papageorgiou, M., METANET: A macroscopic simulation program for motorway networks. *Traffic Engineering and Control* 31, 466-470, 1990.
- [35] Daganzo, C. F. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research*, Vol. 28B, No. 4, pp. 269-287, 1993.
- [36] Daganzo, C. F. The cell transmission model, Part II: Network traffic. *Transportation Research*, Vol. 29B, No. 2, pp. 79-93, 1995.
- [37] Muñoz, J.C., C.F. Daganzo. The bottleneck mechanism of a freeway diverge. *Transportation Research Part A*, 36: 483-508, 2002.
- [38] Jin, W. L., Zhang, H. M. On the Distribution Schemes for Determining Flows through a Merge. *Transportation Research Vol. 37B*, 6, pp. 521-540, 2003.
- [39] Gen, M., R. Cheng, *Genetic Algorithm & Engineering Optimization*, Wiley-IEEE, 2000.
- [40] Wall, M., GALib, a C++ library of genetic algorithm components. Massachusetts Institute of Technology, 1999. <http://lancet.mit.edu/ga/>, Retrieved on 04.16.2009.