An Integrated Control Model for Managing Network Congestion

Heng Hu*
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive S.E.
Minneapolis, MN 55455
Email: huxxx186@umn.edu
(*Corresponding Author)

Henry X. Liu
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive S.E.
Minneapolis, MN 55455
Phone: 612 625 6347
Fax: 612 626 7750
Email: henryliu@umn.edu

Total words: 5282+8 figures + 1 table =7532

Submitted to the
2013 Transportation Research Board Annual Meeting
For both Publication and Presentation
November 15th, 2012
An Integrated Control Model for Managing Network Congestion

Heng Hu and Henry X. Liu
Department of Civil Engineering, University of Minnesota

Abstract
An integrated control model is proposed in this paper to manage traffic congestion along a freeway and a parallel signalized arterial. This model focuses on diversion control, which seeks to utilize available capacity along parallel routes. It specifically considers the potential impact of diverting traffic to the performance of diverting route. For signalized arterial, the caused congestion can be reduced or eliminated by a maximum flow based signal control model. The integrated control model does not need the time-dependent traffic demand information as most of previous approaches do and it is suitable for online applications because of its extremely low computation burden. The model is tested using microscopic traffic simulation in the I-394 and TH 55 corridor in Minneapolis, MN. The results indicate the model can effectively and efficiently reduce network congestion and improve system performance.

Key words: Integrated Corridor Management, Diversion Control, Network Congestion

1. Introduction
Because of the increasing traffic demand and limited facility resources, traffic congestion has become a more and more severe problem for metropolitan areas not only in the United States but also around the world. How to efficiently and effectively manage traffic during peak hours or non-recurrent congestion period appears to be a challenging task for researchers and practitioners. The Integrated Corridor Management (ICM) approach has drawn more and more attention in recent years because it is believed to be a promising tool to manage urban traffic congestion. According to FHWA, the ICM program seeks to “optimize the use of existing infrastructure assets and leverage unused capacity” along urban corridors to help manage congestion.

In fact, researchers have devoted their effort to the ICM domain for decades. In 1988, Van Aerde and Yagar (1988a and 1988b) first clearly stated the importance of integrated control and discussed the required characteristics to operate an integrated control system. Following that, researchers have established various integrated traffic control strategies. The first category of strategies focuses on the provision of route guidance at the bifurcation point through certain media, such as Variable Message Sign (VMS). It was first studied by Papageorgiou (1990), who attempted to conceptually integrate ramp metering, real time information, route guidance and
signal control for freeway corridor management. This approach was extensively studied and
extended by Hawas & Mahmassani (1995), Messmer & Papageorgiou (1995), Ben-Akiva et. al
The decision variables of these models are splitting rates at each bifurcation node. However, since
it is very difficult to estimate drivers’ compliance rates and they vary with time and location, even
if the splitting rates can be optimally calculated, the actual performance would be undermined.
Further, most of these models require a pre-known time-dependent traffic demand as inputs, which
largely limits the practical application of this type of models. Another large category of integrated
traffic control strategies focuses on the interactions between different modes, e.g. freeway and
control approach for traffic corridors including both freeways and signalized arterials based on the
store-and-forward modeling philosophy. Later, Wu and Chang (1999) proposed a control model
which integrates ramp metering, intersection signal timing and off-ramp diversion under
non-recurrent congestion. Liu et al. (2011) introduced a multi-objective optimization model to
maximize the utilization of the available corridor capacity. However, these models are very
complicated and difficult to solve, and even if they can be successfully solved, the ability of these
models to deal with arterial traffic congestion still appears limited.

In practice, the application of the ICM concept is still at the very early stage. Although the U.S.
Department of Transportation’s (USDOT) Intelligent Transportation Systems (ITS) program
launched the ICM Systems initiative in 2005 and eight cities were selected to be the pioneer sites,
most of the work remains at the policy research level, such as cost-benefit analysis, incentive
analysis and agreement analysis. In 1990s, two field implementation projects sponsored by FHWA
were conducted to evaluate the effectiveness of integrated operational strategies for surface street
and freeway systems. The first project was located in the Twin Cities, Minnesota (Sussman, 2000)
and the second was conducted in the City of Irvine, California (MacCarley et al., 2002). The
effectiveness of the ICM strategies has been shown in these field testing projects.

In this paper, a new diversion control model is developed to reduce network congestion by
utilizing available capacity of parallel routes. Comparing with previous control models, the
proposed model has the following merits:

1. The impact of the diversion traffic to diverting route is specifically considered, especially
   for signalized arterial, so the potential congestion caused by diversion traffic is reduced or
   eliminated by proper adjustment of signal timings.
2. The proposed model does not have the requirement on time-dependent traffic demand
   information as model input. It is ready to be implemented at typical parallel traffic
   corridors where the standard detection system is available.
3. The proposed model has very low computation burden and is suitable for on-line
   applications.

In the following, Section 2 defines the problem which will be solved in the paper, followed by the
detailed formulation of the proposed model in Section 3. The case study site and simulation results
are presented in Section 4 and finally the conclusions and remarks of this paper are given in
Section 5.
2. Problem statement

This paper aims to solve the diversion control problem between two alternative routes in order to take advantage of all available capacities. A more typical and challenging situation is that the two routes belong to different control types, e.g. one route is freeway and the other is signalized arterial (see Figure 1). The two origins \( O_1 \) and \( O_2 \) might be the same or different and so do the two destinations \( D_1 \) and \( D_2 \). In practice, most daily commuters would choose one of the routes based on their driving experience and preference with little variation. However, if the performance on one of the routes is significantly worse than the other, which might be caused by either recurrent (e.g. daily congestion during peak hours) or non-recurrent (e.g. car crash) event, diverting a portion of travelers to the alternative route with better performance would certainly benefit the whole system. Considering the diversion control between freeway and signalized arterial, there are actually two embedding sub-problems. One is the control strategy to divert the freeway traffic to the arterial system when freeway has worse performance and the other is the opposite. How to inform travelers with real-time traffic information and how to predict the potential impacts of diverting traffic to the diverting route are the two most important questions which need to be answered in this paper.

![Figure 1 Problem statement](image)

3. Model formulation

3.1 Performance estimation

In order to make correct control decisions, the performance of both routes needs to be monitored in real-time. At the end of each control period \( t \), control decisions for the next control period \( t+1 \) will be made based on the traffic conditions in the immediate past control period \( t \). The control interval usually includes 2–3 signal cycles. In this sub-section, the performance estimation method for both freeway and signalized arterial will be introduced.

3.1.1 Freeway performance estimation
Density and travel time are the two most important measures to reflect freeway performance. To estimate the real-time density and travel time on freeway, certain detection system (e.g. loop detectors, cameras, blue tooth technology and etc) is assumed to be available. Loop detector is one of the most commonly used techniques in the current traffic infrastructure. Detector stations are usually placed every 0.5 to 1 mile along freeways. The loop detector data, such as volume, density and speed, is transferred back to the control center in aggregated levels (e.g. every 30 seconds or 1 minute). In the proposed control model, a freeway corridor is divided into segments such that each segment contains at least one detector station. The performance of each segment is estimated based on the collected data from corresponding detector station. Assume the freeway is divided into $M$ segments, $1, 2, ..., M$. The density of each segment is denoted by $k_m(t)$ (Vehicles/Mile) and the average speed of each segment is $v_m(t)$ (Miles/Hour) at control period $t$. Thus, the travel time along the freeway corridor, denoted by $T^f(t)$, can be calculated by (1), where $l_m$ is the length for segment $m$. Based on historical traffic data, it is not difficult to find out the critical density value $k^c_m$ (i.e. the density when flow reaches maximum) for each segment. At the control period, it is desired that the density of each freeway segment will always stay below the corresponding critical value $k^c_m$, otherwise severe congestion might be introduced to the freeway system. Freeway residual capacity $\eta^f(t)$ is defined as (2). If $\eta^f(t) > 0$, extra traffic demand can be handled by the freeway corridor; If $\eta^f(t) \leq 0$, the freeway corridor is saturated or over-saturated.

$$ T^f(t) = \sum_{m=1,2,...,M} \left[ l_m / v_m(t) \right] $$  \hspace{1cm} (1)

$$ \eta^f(t) = \min_{m=1,2,...,M} \left[ k^c_m - k_m(t) \right] $$  \hspace{1cm} (2)

### 3.1.2 Arterial performance estimation

In order to estimate the arterial performance in real-time, the arterial traffic data collection system is also expected to be available, for instance, the SMART-SIGNAL system (Liu, et al., 2009), which automatically archives the event-based high-resolution traffic data (i.e. signal changes and vehicle actuations). Based on the collected data set, real-time queue length can be estimated from cycle to cycle with very high accuracy (Liu, et al., 2009). Further, the arterial travel time $T^a(t)$ at control period $t$ can be estimated by the virtual probe approach proposed in Liu and Ma (2009). The algorithm mimics the behaviors (e.g., acceleration, deceleration, no speed change and etc.) of a virtual vehicle traveling from origin to destination based on real-time traffic information (i.e. signal and queue) and the accuracy has been verified in several field projects.
Before making any adjustment, the residual capacity of each intersection needs to be calculated. Assume that there are $N$ intersections along the signalized arterial, the residual capacity of intersection $n$ (denoted by $\eta_{n,i}(t)$) for the phase of diverting traffic direction (i.e. phase $i$) during the control period $t$ can be calculated by (3). $g_{n,i}(t)$ is the green time for phase $i$ of intersection $n$ during control period $t$, $s_{n,i}$ is the corresponding saturation flow rate, $z_{n,i}$ is number of lanes and $\gamma_{n,i}(t)$ is the average cycle discharging volume for phase $i$ of intersection $n$ during control period $t$. The residual capacity measures how much more traffic can be discharged during one cycle at specific intersection based on the current traffic condition.

$$\eta_{n,i}(t) = g_{n,i}(t)s_{n,i}z_{n,i} - \gamma_{n,i}(t), n \in \{1, 2, ..., N\}$$  

(3)

The residual capacity along the signalized arterial is the minimum residual capacity among all intersections,

$$\eta^*(t) = \min_{n \in \{1, 2, ..., N\}} \left[ \eta_{n,i}(t) \right]$$  

(4)

When traffic along signalized arterial becomes congested, oversaturated traffic conditions may happen, which will cause detrimental effects to signal operation. An Oversaturation Severity Index (OSI) was proposed by Wu et al. (2010) to quantify the severity level of oversaturation by measuring its detrimental effects. Detrimental effect is characterized by either a residual queue at the end of a cycle or a spillover from downstream traffic, both of which create “unusable” green time. In the case of residual queue, the “unusable” green time is the equivalent green time to discharge the residual queue in the following cycle, but for spillover, the “unusable” green time is the time period during which an downstream link is blocked therefore the discharge rate is zero. OSI is further differentiated into TOSI (Temporal Oversaturation Severity Index, caused by the residual queue that creates the detrimental effect in temporal dimension) and SOSI (Spatial Oversaturation Severity Index, caused by the spillover that creates the detrimental effect in spatial dimension). In the following, we use $S_{n,i}(t)$ to represent the unusable green time caused by spillover at intersection $n$ phase $i$ during time period of $t$, and use $T_{n,i}(t)$ to represent the unusable green time caused by residual queue.

### 3.2 Diversion Control

Diversion control decisions are made based on the real-time estimated performance on both routes.
3.2.1 From arterial to freeway

When the performance on signalized arterial is worse than that on freeway at certain control period $t$, we may want to divert some traffic from arterial to freeway to reach better system performance. To be more specific, if the travel time on the arterial ($T^a(t)$) is longer than the travel time on the freeway ($T^f(t)$) plus the diversion cost (e.g. $T^{a\rightarrow f}(t)$, travel time on diverting links), see (5), the diversion control from arterial to freeway is preferred.

$$T^a(t) > T^{a\rightarrow f}(t) + T^f(t) \quad (5)$$

To warrant the diversion, we need to further check the freeway residual capacity $\eta^f(t)$. If $\eta^f(t) > 0$, the diversion control from arterial to freeway is warranted; otherwise, the diversion control is declined. If the diversion control is warranted, a variable message sign (VMS) can be shown on the arterial side before the diverting point, advising drivers to use the freeway system, see Figure 2. During the whole control period, the density of each segment on freeway should always be kept lower than the critical density $k_m^c$, otherwise severe congestion might be introduced to the freeway system. To make sure this, the diverting traffic volume from arterial to freeway in the next control period $t+1$, denoted by $\pi^{a\rightarrow f}(t+1)$, should be smaller than the allowable traffic volume increase on freeway, as presented in (6), where $v^f$ is the free flow speed on freeway and $\Delta t$ is the control interval. The restriction on $\pi^{a\rightarrow f}(t+1)$ can be achieved by the signal control at the diverting intersection.

$$\pi^{a\rightarrow f}(t+1) \leq \eta^f(t)v^f\Delta t \quad (6)$$

![Figure 2 Diversion control from arterial to freeway](image-url)
3.2.2 From freeway to arterial

On the other hand, when the performance on freeway is worse than that on arterial at certain control period \( t \), we may want to divert some traffic from freeway to arterial, see Figure 3. This condition can be expressed by (7), where \( T^{f \rightarrow a}(t) \) is the diversion cost from freeway to arterial, i.e. travel time on diverting links. In this case, a variable message sign (VMS) can be shown on the freeway side before the diverting point, indicating the travel times on both routes.

\[
T^{f}(t) > T^{f \rightarrow a}(t) + T^{a}(t)
\]  

(7)

In this case, since there is usually no signal control on the freeway mainline, the exact number of diverting traffic is difficult to control, although the VMS can give drivers advice. To overcome this problem, a Logit decision model is used to predict the diversion rate \( \phi(t+1) \) at the next control period \( t+1 \), i.e. the percentage of vehicles that will be diverted from freeway to arterial because of the provision of traffic information on VMS. In the model, the diversion rate \( \phi(t+1) \) is calculated based on travel time difference between arterial and freeway as shown in (8), where \( u \) is the travel time difference in minutes, \( \beta \) is the coefficient that values travel time with respect to travel utility, and \( \alpha \) is the parameter that represents every other factor not related to time, such as drivers’ inertia of mind (i.e. unwillingness to divert). Both \( \alpha \) and \( \beta \) can be estimated based on historical data and experience.

\[
\phi(t+1) = \frac{1}{1 + e^{\alpha + \beta u(t)}}
\]  

(8)

\[
u(t) = T^{f}(t) - [T^{a}(t) + T^{f \rightarrow a}(t)]
\]

The diverted traffic volume from freeway to arterial in the next control period \( t+1 \), denoted by \( \pi^{f \rightarrow a}(t+1) \), can be predicted by (9) based on the assumption that the incoming traffic during control period \( t+1 \) at the freeway is the same as that during control period \( t \), denoted by \( q(t) \).

Because of the diversion traffic into the arterial system, the signal timings along arterial need to be adjusted accordingly.

\[
\pi^{f \rightarrow a}(t+1) = q(t) \phi(t+1) \Delta t
\]  

(9)
If \( c \times \pi^{f-a}(t+1)/\Delta t \leq \eta^a(t) \), the diverting traffic can be handled by the current signal timings along the arterial; however, if \( c \times \pi^{f-a}(t+1)/\Delta t > \eta^a(t) \), the diverting traffic will cause residual queue at some intersection(s). The residual queue \( \psi_{n,d}(t+1) \) at each intersection during the next control period \( t+1 \) can be predicted by (10), where \( \Delta \lambda_{n,d}(t+1) \) is the predicted increase of arrival traffic at intersection \( n \) during control period \( t+1 \). The initial condition is \( \Delta \lambda_{1,j}(t+1) = c \times \pi^{f-a}(t+1)/\Delta t \). The first equation basically says if the increase of arrival flow at specific intersection during control period \( t+1 \) (i.e. \( \Delta \lambda_{n,j}(t+1) \)) is larger than the corresponding residual capacity (i.e., \( \eta_{n,d}^a(t) \)), there will be a residual queue \( \Delta \lambda_{n,j}(t+1) - \eta_{n,d}^a(t) \); otherwise, there will be no residual queue at intersection \( n \). The second equation updates the increase of arrival flow to the downstream intersection (i.e. \( \Delta \lambda_{n+1,j}(t+1) \)), which is equal to the minimum of the residual capacity (i.e., \( \eta_{n,d}^a(t) \)) and the increase of arrival flow (i.e. \( \Delta \lambda_{n,j}(t+1) \)) at the current intersection.

\[
\begin{align*}
\psi_{n,j}(t+1) &= \max\left[0, \Delta \lambda_{n,j}(t+1) - \eta_{n,d}^a(t)\right] \\
\Delta \lambda_{n+1,j}(t+1) &= \min\left[\eta_{n,d}^a(t), \Delta \lambda_{n,j}(t+1)\right], n \in \{1, 2, ..., N\}
\end{align*}
\]

When residual queue happens at signalized intersections, it means the current discharging capacity cannot accommodate the increase of traffic. If the signal timings are not properly adjusted, more severe oversaturated conditions, such as spillovers, will appear. Therefore, a systematic approach is needed to mitigate or eliminate oversaturated traffic conditions between intersections and it will
be introduced in the next section.

3.3 A maximum flow based signal control model

Based on the predicted potential impact to the arterial (i.e. $\psi_{n,i}(t+1)$) and the real-time estimated arterial oversaturation level (i.e. $S_{n,i}(t)$ and $T_{n,i}(t)$), a maximum flow based signal control model is developed in this section.

3.3.1 Control Variables

Along the traffic direction, two control variables $\Delta r_{n,i}(t)$ and $\Delta g_{n,i}(t)$, namely red time changes and green time changes for phase $i$ of intersection $n$, are introduced for each signalized intersection. Whether to change red or green is determined by the causes of the oversaturation. Changing red times (i.e. $\Delta r_{n,i}$) aims to eliminate spillover; and changing green times (i.e. $\Delta g_{n,i}$) aims to clear residual queues. A positive red time change (red extension) means that extra red time is added. Since the cycle length is kept unchanged, the green start would be postponed with the red extension (see Figure 4 a) and the total green time is reduced. A negative red time change (red reduction) means a portion of red time is cut from the end of red, therefore, green start will be advanced (see Figure 4 b) and the total green time is increased. Similarly, a positive green time change (green extension) indicates that additional green time is added to the original end of the green time (see Figure 4 c), and a negative green time change (green reduction) represents that some green time is cut from the end of green (see Figure 4 d). Depending on the offset reference point used for the intersection (start of yellow, start of green, barrier crossing, etc.), each case of adjusting green or red may require a corresponding change to the offset and green split values.

![Figure 4](image-url)  
**Figure 4**  Red Time Changes & Green Time Changes

3.3.2 Constraint Analysis

(1) Spillover Elimination

The proposed control model aims to eliminate spillover between intersections. In order to
eliminate the spillover at intersection \( n \), one can either extend the red time at the current intersection \( n \) (i.e., apply gating at the upstream intersection), or reduce the red time at the downstream intersection \( n+1 \) (i.e., discharge the downstream queue earlier), or a combination of the two strategies. As described in Figure 5, extending the red at intersection \( n \) by \( \Delta r_{n,j}(t+1) \) \((\Delta r_{n,j}(t+1) > 0) \) will make the unusable green time caused by spillover shorter by \( \Delta r_{n,j}(t+1) \); On the other hand, reducing the red at intersection \( n+1 \) by \( \Delta r_{n+1,j}(t+1) \) \((\Delta r_{n+1,j}(t+1) < 0) \) will make the unusable green caused by spillover at intersection \( n \) shorter by \( \Delta r_{n+1,j}(t+1) \). Therefore, in order to eliminate spillover at intersection \( n \), the difference of red time changes between intersection \( n \) and intersection \( n+1 \) should be equal to the unusable green time caused by spillover at intersection \( n \), i.e., \( S_{n,j}(t) \), see Eq. (11).

\[
\Delta r_{n,j}(t+1) - \Delta r_{n+1,j}(t+1) = S_{n,j}(t), \forall n \in \{1, ..., N-1\} \tag{11}
\]

(2) Residual Queue Elimination

If Eq.(11) is satisfied, the spillovers are supposed to be eliminated during control period \( t+1 \). Then the green time change \( \Delta g_{n,j}(t+1) \) for each intersection is used to eliminate residual queue. If the red time and green time changes at intersection \( n \) are \( \Delta r_{n,j}(t+1) \) and \( \Delta g_{n,j}(t+1) \) respectively, the total green time at intersection \( n \) for control period \( t+1 \) would be \([\Delta r_{n,j}(t+1) + \Delta g_{n,j}(t+1) + \bar{g}_{n,j}(t)]\). If Intersection \( n+1 \) has residual queue in control period \( t \) and the corresponding unusable green time is \( T_{n+1,j}(t) \), in order to eliminate residual queue of Intersection \( n+1 \) at control period \( t+1 \), the difference of total green time between Intersection \( n+1 \) and its upstream intersection \( n \) should be equal to \( T_{n+1,j}(t) \). However, because

![Figure 5 An example of applying strategies 2 & 3 to eliminate spillover](image)
of the diversion control, extra residual queue $\psi_{n_{t+1},i}(t+1)$ will be introduced at intersection $n+1$ at the next control period $t+1$, considering that, Eq. (12) should hold.

$$
\left[ \Delta g_{n_{t+1},i}(t+1) - \Delta r_{n_{t+1},i}(t+1) + g_{n_{t+1},i}(t) \right] - \left[ \Delta g_{n_i}(t+1) - \Delta r_{n_i}(t+1) + g_{n_i}(t) \right] = T_{n_{t+1},i}(t) + \psi_{n_{t+1},i}(t+1) / \left( z_{n_{t+1},i} s_{n_{t+1},i} \right), n \in \{1, \ldots, N-1\}
$$

(12)

Substitute (11) into (12),

$$
\Delta g_{n_{t+1},i}(t+1) - \Delta g_{n_i}(t+1) = T_{n_{t+1},i}(t) + \psi_{n_{t+1},i}(t+1) / \left( z_{n_{t+1},i} s_{n_{t+1},i} \right) - S_{n_{t+1},i}(t) - \left[ g_{n_{t+1},i}(t) - g_{n_i}(t) \right], n \in \{1, \ldots, N-1\}
$$

(13)

(3) Available Green Constraints

For each intersection along the oversaturated route, the green time increase at control period $t+1$, i.e., $\Delta g_{n,i}(t+1) - \Delta r_{n,i}(t+1)$ is constrained by the available green time $g_{n,i}^a(t+1)$ for intersection $n$ and phase $i$, see (14).

$$
\Delta g_{n,i}(t+1) - \Delta r_{n,i}(t+1) \leq g_{n,i}^a(t+1), \forall n \in \{1, \ldots, N\}
$$

(14)

If $Z_{n,i}$ is the set of conflicting phases to phase $i$ at intersection $n$, the available green time $g_{n,i}^a(t+1)$ can be computed by considering the maximum queue size for each of these conflicting phases in the immediate past control interval $t$, see Eq.(15). Here $c_n$ is the cycle length for intersection $n$, $q_{n,p}^{max}(t)$ is the maximum queue size per lane for phase $p$ at intersection $n$ at control interval $t$ and $s_{n,p}$ is the saturation flow rate per lane for phase $p$ of intersection $n$.

$$
q_{n,p}^{max}(t) / s_{n,p} \text{ calculates how much green time is needed to discharge the queue of } q_{n,p}^{max}(t). a \text{ is a weighting term, which represents users’ perspective on the importance of queues on conflicting phases when calculating the available green for oversaturated phase $i$. When the maximum queue length for phase $p$, i.e. } q_{n,p}^{max}(t)h \text{ (where } h \text{ is the jammed space headway), is shorter than the corresponding link length } L_{n,p}, \text{ we would like to only account for a portion } (a = b, b \in (0,1)) \text{ of these queues because we want to maximize the discharging capacity for the oversaturated route to reduce congestion; however, if the maximum queue length for phase $p$ is already longer than the link length, all the queues need to be considered } (a = 1), \text{ otherwise these queues will block further upstream intersections. One should note that, the smaller the $b$ is, the more expected extra capacity will be assigned to the oversaturated route, the faster the queues on conflicting phases
will grow and the more delay will be introduced to conflicting phases. A recommended value for $b$ would be around 0.5.

$$g_{n,j}(t+1) = c_n - \sum_{p \in Z_{n,j}} a \left[ \frac{q_{n,p}^{\max}(t)}{s_{n,p}} \right] - g_{n,j}(t) \quad (15)$$

Where $a = \begin{cases} b, & \text{if } q_{n,p}^{\max}(t)h < L_{n,p} \\ 1, & \text{if } q_{n,p}^{\max}(t)h \geq L_{n,p} \end{cases}$

### 3.3.3 Signal control model

The objective of the control model is to maximize the discharging capacity along the oversaturated route. At each control period $t$, it is equivalent to maximizing the total green time at the first intersection of the route, i.e., \( (\Delta g_{i,j}(t+1) - \Delta r_{i,j}(t+1) + g_{i,j}(t)) \). Since $g_{i,j}(t)$ is the green time during control period $t$, at the start of control period $t+1$, maximizing $\left( \Delta g_{i,j}(t+1) - \Delta r_{i,j}(t+1) + g_{i,j}(t) \right)$ is equivalent to maximizing $\left( \Delta g_{i,j}(t+1) - \Delta r_{i,j}(t+1) \right)$.

Therefore, the complete control model can be expressed in (16). The first and second constraints ensure the elimination of spillover and residual queues between intersections and the third constraint considers the available green time. All the constraints and objective function of the model are linear, so it can be solved with very little computation burden, which makes it suitable for online applications.

$$\begin{align*}
\max & \quad \Delta g_{i,j}(t+1) - \Delta r_{i,j}(t+1) \\
\text{s.t.,} & \\
\Delta r_{n,j}(t+1) - \Delta r_{n+1,j}(t+1) &= S_{n,j}(t), \quad n \in \{1, \ldots, N-1\} \\
\Delta g_{n,j}(t+1) - \Delta g_{n+1,j}(t+1) &= T_{n,j}(t) + \psi_{n+1,j}(t+1)/\left( z_{n+1,j}s_{n+1,j} \right) - S_{n,j}(t) - \left[ g_{n+1,j}(t) - g_{n,j}(t) \right], \quad n \in \{1, \ldots, N-1\} \\
\Delta g_{n,j}(t+1) - \Delta r_{n,j}(t+1) &\leq g_{n,j}^{a}(t+1), \quad n \in \{1, \ldots, N\} \\
g_{n,j}(t+1) &= c_n - \sum_{p \in Z_{n,j}} a \left[ \frac{q_{n,p}^{\max}(t)}{s_{n,p}} \right] - g_{n,j}(t), \quad n \in \{1, \ldots, N\}
\end{align*}$$

(16)

### 4. Case study and simulation

In order to test the proposed approach, a case study site was selected in Minneapolis, MN. As shown in Figure 6, there are two major routes, i.e. Trunk Highway 55 (a coordinated high speed signalized arterial) and Interstate freeway 394, connecting the west suburban living areas and the downtown Minneapolis. The total length of the corridor is about 3.5 miles and both routes (i.e. I-394 and TH 55) have a speed limit of 55 MPH. The coordination of the TH 55 favors the eastbound traffic during the AM peak hours because of the large traffic from home to work and it
favors the westbound during the PM peak hours because of the returning traffic. Based on the detector station locations in the field, the I-394 freeway is divided into 6 segments (see Figure 6) such that each segment contains one detector station. Figure 7 shows the flow-density diagram from the three detectors (one for each lane) at segment 4 based on the field collected data between 6/15/2009 and 6/19/2009. One can easily find out that the critical density for segment 4 is about 150 Vehicles/Mile/Lane.

Figure 6 Case study site: the TH 55/I-394 corridor, Minneapolis, MN

Figure 7 Flow-density diagram from three detectors at segment 4

A VISSIM model is then built and calibrated using the field data collected during the morning peak hours (7:00 AM – 9:00 AM) between 6/15/2009 and 6/19/2009, see Figure 8. Because of the space limit, only the diversion control from freeway to arterial is tested in this paper. The diverting route is shown in Figure 8 by the green dotted line, which goes through TH 169 northbound, TH 55 eastbound and then TH 100 southbound. The simulation control program was written in C# and it controls the simulation in real-time through the COM interface of VISSIM. At each control period, the travel time on freeway (i.e. $T^f(t)$) is estimated based on the segment speed; the travel time on the arterial (i.e. $T^a(t)$) is estimated through the virtual probe approach; the diverting cost $T^{f\rightarrow a}(t)$ is the summation of travel times on TH 169 northbound and TH 100 southbound, which can also be estimated through the same approach as freeway.
The simulation lasts for two hours (7:00 AM – 9:00 AM) and Figure 9 shows the demand profiles of the major directions (i.e. I-394 EB, I-394 WB, TH 55 EB and TH 55 WB) for the whole simulation period. The cycle length of the signalized arterial is 180 seconds and the control interval $\Delta t$ is 360 seconds. To simulate some unexpected incident (i.e. car crash) happening on freeway, a reduced speed area (10 MPH) with a length of 800 ft is created on the eastbound of I-394 from 7:30 AM to 8:30 AM (see Figure 8). Vehicles passing that area during that time window have to reduce their speed and as a result severe congestion will happen on the eastbound of I-394. In the following, two scenarios will be tested: one is the base scenario with original control strategy (i.e. independent control) and the other is the scenario with the proposed integrated control strategy. Each scenario is run for 10 times using different random seeds and the average results are listed below.

Table 1 summarizes the network performance during the whole simulation period. With original
control strategy, the average delay is 55.69 Seconds/Veh, while with the proposed integrated control strategy, the average delay is reduced to 41.14 Seconds/Veh, which is a 26.13% reduction. For average number of stops of the whole network, the proposed control model reduces it from 2.21 to 1.28, a 42.13% reduction. The average speed is increased from 42.12 MPH to 45.86 MPH.

Table 1. Network performance comparison

<table>
<thead>
<tr>
<th></th>
<th>Base Scenario</th>
<th>With diversion</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Delay (Seconds per veh.)</strong></td>
<td>55.69</td>
<td>41.14</td>
<td>-26.13%</td>
</tr>
<tr>
<td><strong>Average # of stops (per veh.)</strong></td>
<td>2.21</td>
<td>1.28</td>
<td>-42.13%</td>
</tr>
<tr>
<td><strong>Average Speed (MPH)</strong></td>
<td>42.12</td>
<td>45.86</td>
<td>+8.89%</td>
</tr>
</tbody>
</table>

In order to test the performance of the proposed control strategy to handle different demand levels, we increase and decrease the mainline demand (i.e. the demand shown in Figure 9) by 5% and then run the simulation again. Table 2 presents the network performance under demand variations. When the mainline demand is increased by 5%, the whole network becomes more congested, which can be reflected by the increase of average delay and average number of stops and the decrease of average speed. However, with the proposed diversion control strategy, average delay and average number of stops can be reduced by 16.31% and 38.2% respectively and average speed can be increased by 7%. On the other hand, when the mainline demand is decreased by 5%, the proposed diversion control strategy can still significantly improve the network performance, i.e. reduce average delay by 29.67%, reduce average number of stops by 47.97% and increase average speed by 7.82%. Based on the results discussed above, one can see that the proposed integrated control model can effectively reduce network congestion and smooth traffic movement by utilizing the available capacity along parallel route.

Table 2. Network performance comparison with demand variations

<table>
<thead>
<tr>
<th></th>
<th>Increase demand by 5%</th>
<th>Decrease demand by 5%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base Scenario</td>
<td>With diversion</td>
</tr>
<tr>
<td><strong>Average Delay (Seconds per veh.)</strong></td>
<td>76.79</td>
<td>64.27</td>
</tr>
<tr>
<td><strong>Average # of stops (per veh.)</strong></td>
<td>3.45</td>
<td>2.13</td>
</tr>
<tr>
<td><strong>Average Speed (MPH)</strong></td>
<td>37.63</td>
<td>40.26</td>
</tr>
</tbody>
</table>

5. Conclusion
In this paper, we propose an integrated control model to manage network congestion. Through diversion control, the model tries to fully utilize the available capacity along parallel routes. The impact of the diversion traffic is specifically considered, especially for signalized arterial, so caused congestion can be reduced or eliminated by proper adjustment of signal timings. This model does not rely on time-dependent traffic demand as model inputs. It is ready to be implemented at typical parallel traffic corridors where the standard detection system is available. With the extremely low computation burden and the model is very suitable for on-line applications. We have tested the performance of the proposed model using microscopic traffic simulation in the I-394 and TH 55 corridor in Minneapolis, MN. The results indicate that the proposed model significantly reduces the network congestion and makes traffic much smoother, which can be reflected by the huge improvement on network performance measures, such as average delay per vehicle, average number of stops per vehicle and average speed. For future research, the freeway ramp metering control should be included in the model formulation. More scenarios should be tested not only in the simulation but hopefully in real field.

References:


