

**A SET OF NEW TRAFFIC-RESPONSIVE
RAMP-METERING ALGORITHMS AND
MICROSCOPIC SIMULATION RESULTS**

Xiaotian Sun

Department of Mechanical Engineering
University of California, Berkeley
Berkeley, CA 94720-1742
Tel: 510-643-8903, Fax: 510-643-5599
E-mail: sunx@berkeley.edu

Roberto Horowitz

Department of Mechanical Engineering
University of California, Berkeley
Berkeley, CA 94720-1742
Tel: 510-642-4675, Fax: 510-643-5599
E-mail: horowitz@berkeley.edu

Transportation Research Board
The 85th Annual Meeting
January 22–26, 2006
Washington, D.C.

November 15, 2005

Text: 4752 words

9 figures: $250 \times 9 = 2250$ words

1 table: $250 \times 1 = 250$ words

Total: 7252 words

ABSTRACT

In this paper, we present a novel switching traffic-responsive ramp-metering controller that adapts to the different traffic dynamics under different congestion conditions—free-flow or congested. The approach of multirate linear quadratic control with integral action (LQI) is employed to compensate for disturbances and to accommodate the difference between the model sampling time and the metering-rate update interval. In addition, a queue length regulator is designed to prevent the queue from exceeding the ramp storage capacity and yield improved performance over the currently used *ad hoc* “queue-override” scheme. Subsequently, a queue length estimator is designed to provide feedback to the queue length regulator, using the queue-detector speed data that are available in the field. A local ramp-metering control strategy is proposed to achieve the control goal of reducing the spatial and temporal span of the congestion, while satisfying the on-ramp storage capacity constraints, using locally available information. Test results on a calibrated microscopic traffic simulator demonstrate the performance and effectiveness of the switching ramp-metering controller, the queue length estimator and regulator, and the overall control strategy. The Total Vehicle and Passenger Congestion Delays are both reduced by 16%, while the Total Travel Time is improved by 5.6%. As a comparison, simulation results of ALINEA are also presented.

1 INTRODUCTION

Freeway traffic congestion is a major problem in today's metropolitan areas. It occurs regularly during commute hours. In addition, non-recurrent congestion often takes place as a result of incidents, road work, or public events. Congestion causes inefficient operation of freeways, wasting of resources, increased pollution, and intensified driver fatigue.

The 2004 Urban Mobility Report (1) finds: "Congestion has grown everywhere in areas of all sizes. Congestion occurs during longer portions of the day and delays more travelers and goods than ever before." In the report, the authors have calculated that in 2002, congestion cost Americans 3.5 billion hours of delay and 5.7 billion gallons of wasted fuel, with an equivalent monetary cost of \$63.2 billion.

On-ramp metering has been widely used as an effective strategy to increase freeway operation efficiency. It has been recommended to the Federal Highway Administration as the No. 1 tool to address the congestion problem, other than adding more capacity to transportation infrastructures (2). It has been reported that ramp metering was able to reduce delay by 101 million person-hours in 2002, approximately 5% of the congestion delay on freeways where ramp-metering was in effect (1).

In this paper, we present a novel switching ramp-metering controller that employs a different feedback structure depending on whether the freeway is in a *free-flow* or *congested* mode. It has been known that the traffic dynamics behave differently under free-flow or congested conditions. This leads to different controllability and observability properties for free-flow versus congested traffic (3). It is thus natural to design different ramp-metering controllers for these different congestion modes.

In addition to the switching mainline traffic responsive ramp-metering controller, we also present a proportional-integral (PI) regulator to keep the on-ramp queue below the storage capacity limit and prevent excessively long queues from interfering with surface street traffic. This PI regulator will have better performance than the "queue-override" scheme currently used on California freeways. We also design and validate a queue length estimator using the speed data measured by the queue detector. This estimator provides the feedback that is needed by the queue length regulator.

We also propose a localized control strategy for the switching metering controller and the queue length regulator to achieve the goal of reducing the spatial and temporal span of the congestion.

Test results of this set of ramp-metering algorithms, as well as those of ALINEA (4), on a calibrated microscopic traffic simulator are also presented.

2 PERFORMANCE MEASURES

In this section, some performance measures are defined for quantitative evaluation of a given freeway segment. All the quantities are defined for the time period T and the freeway segment L .

TTD_V Total Travel Distance, which is defined as the sum of the distances traveled by all the vehicles in L within T .

TTT_V Total Travel Time, which is the sum of the time that is spent by all vehicles in L within T . It includes the time spent by vehicles waiting in the on-ramp queues.

TCD_V Total Congestion Delay, which is the difference between the Total Travel Time and the time that would be spent by all the vehicles if there were no congestion. $TCD_V = TTT_V - TTD_V/v_0$, where v_0 is the nominal free-flow speed.

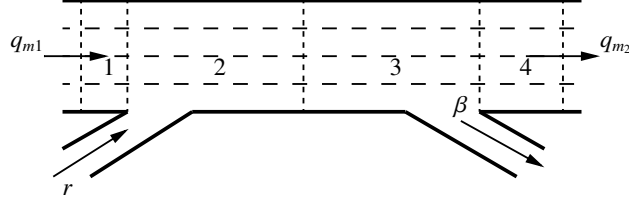


FIGURE 1 A schematic plot of a 4-cell freeway section with one on-ramp and one off-ramp.

Another set of passenger-weighted performance measures can be defined by first collecting the traffic quantities separately for the low- and high-occupancy vehicle classes, and then weighting these quantities by the average passenger number in each vehicle class when calculating the performance measures. This set of passenger-weighted performance measures include Total Passenger Travel Distance TTD_p , Total Passenger Travel Time TTT_p , Total Passenger Congestion Delay TCD_p .

3 SWITCHING-MODE MODEL AND CONTROLLABILITY

In our previous work (3), we piecewise linearized the cell transmission model (CTM) (5, 6) and derived a switching-mode traffic model. Both this switching-mode model and the CTM have been calibrated and tested (3, 7) using the actual traffic data collected from a segment of I-210. The switching-mode model was able to reproduce the traffic behavior as accurately as the CTM.

Depending on the traffic conditions in a freeway section, such as the one shown in Fig. 1, the traffic is in a different mode, e.g., *free-flow* or *congested*. In each mode, the vehicle densities in the cells, denoted by 1–4 in the figure, evolve according to a different set of difference equations.

In *free-flow* mode,

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (t+1) = \begin{bmatrix} 1 - \frac{v_{f1}T_s}{l_1} & 0 & 0 & 0 \\ \frac{v_{f1}T_s}{l_2} & 1 - \frac{v_{f2}T_s}{l_2} & 0 & 0 \\ 0 & \frac{v_{f2}T_s}{l_3} & 1 - \frac{v_{f3}T_s}{l_3} & 0 \\ 0 & 0 & (1-\beta)\frac{v_{f3}T_s}{l_4} & 1 - \frac{v_{f4}T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (t) + \begin{bmatrix} 0 \\ \frac{T_s}{l_2} \\ 0 \\ 0 \end{bmatrix} r(t) + \begin{bmatrix} \frac{T_s}{l_1} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} q_{m1} \\ q_{m2} \end{bmatrix} (t) \quad (1)$$

$$= A(1)\rho(t) + B_r(1)r(t) + B_m(1)q_m(t), \quad (2)$$

where ρ_i is the vehicle density in cell i , q_{m1} and q_{m2} are the mainline entering and exiting flows, respectively, r is the on-ramp flow, β is the split ratio of the off-ramp flow, l_i is the length of cell i , and T_s is the sampling time.

In *congested* mode,

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (t+1) = \begin{bmatrix} 1 - \frac{w_{c1}T_s}{l_1} & \frac{w_{c2}T_s}{l_1} & 0 & 0 \\ 0 & 1 - \frac{w_{c2}T_s}{l_2} & \frac{w_{c3}T_s}{l_2} & 0 \\ 0 & 0 & 1 - \frac{w_{c3}T_s}{l_3} & \frac{1}{1-\beta}\frac{w_{c4}T_s}{l_3} \\ 0 & 0 & 0 & 1 - \frac{w_{c4}T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} (t) + \begin{bmatrix} \frac{T_s}{l_1} \\ 0 \\ 0 \\ 0 \end{bmatrix} r(t) + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -\frac{T_s}{l_4} \end{bmatrix} \begin{bmatrix} q_{m1} \\ q_{m2} \end{bmatrix} (t) \quad (3)$$

$$+ \begin{bmatrix} \frac{w_{c1}T_s}{l_1} & -\frac{w_{c2}T_s}{l_1} & 0 & 0 \\ 0 & \frac{w_{c2}T_s}{l_2} & -\frac{w_{c3}T_s}{l_2} & 0 \\ 0 & 0 & \frac{w_{c3}T_s}{l_3} & -\frac{1}{1-\beta}\frac{w_{c4}T_s}{l_3} \\ 0 & 0 & 0 & \frac{w_{c4}T_s}{l_4} \end{bmatrix} \begin{bmatrix} \rho_{J1} \\ \rho_{J2} \\ \rho_{J3} \\ \rho_{J4} \end{bmatrix},$$

$$= A(2)\rho(t) + B_r(2)r(t) + B_m(2)q_m(t) + B_J(2)\rho_J, \quad (4)$$

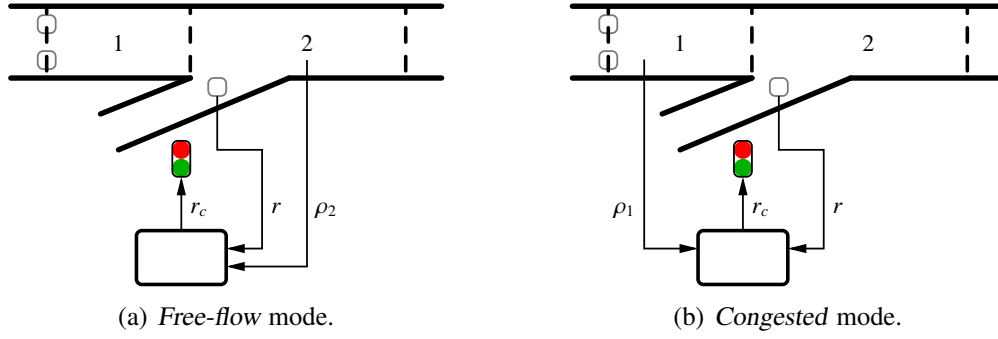


FIGURE 2 Different control structures for different congestion modes.

where ρ_{ji} is the jam density (maximum allowable density) in cell i .

It should be noted that the traffic dynamics are very much different in different modes, which is evidenced by the structures of the “A” matrices in (1) and (3): in *free-flow* mode, the “A” matrix is lower bi-diagonal, and thus the vehicle densities in the downstream cells are affected by those in the upstream cells, i.e., the information travels from upstream to downstream; in *congested* mode, the “A” matrix is upper bi-diagonal, and the vehicle densities in the upstream cells are affected by those in the downstream cells, i.e., the information travels from downstream to upstream.

This observation is critical to the design of an on-ramp metering algorithm, because it determines a dynamical system’s fundamental properties of controllability and observability. A simple calculation reveals that when a freeway section is in *free-flow* mode, the on-ramp can control the vehicle densities downstream, while in *congested* mode, the on-ramp can control the vehicle densities upstream. Therefore, in a different mode, a different feedback structure has to be employed for the metering controller, as shown in Fig. 2.

Although the congestion modes or the cell vehicle densities are not observed or measured directly, we have designed and implemented a mixture Kalman filter (MKF) based estimator (8, 9) that is able to accurately estimate these quantities in real time. The mainline vehicle flows and densities, as well as the on-ramp flows, measured at the locations where loop detectors are installed, are used as input and feedback to the Kalman filters. The estimated congestion mode is used to determine the appropriate control structure, and the estimated vehicle densities are used as feedback.

4 MULTIRATE LQI DESIGN

To compensate for disturbances and to accommodate the difference between the model sampling time and the metering-rate update interval, a multirate linear quadratic control with integral action (multirate LQI) approach (10) was used to synthesize the ramp-metering controller for both of the congestion modes.

In either of the congestion modes, the difference equations governing the evolution of the cell vehicle densities can be written as

$$\rho(t+1) = A\rho(t) + B_r r(t) + B_m q_m(t) + B_J \rho_J, \quad (5)$$

where $r(t)$ is the on-ramp flow, $q_m(t)$ are the mainline flows, and ρ_J are the jam densities.

The desired densities $\bar{\rho}$ have to satisfy the steady-state condition

$$\bar{\rho} = A\bar{\rho} + B_r\bar{r} + B_m\bar{q}_m + B_J\rho_J, \quad (6)$$

where \bar{q}_m is the nominal mainline flow. The error dynamics is thus given by

$$\tilde{\rho}(t+1) = A\tilde{\rho}(t) + B_r r(t) - B_r\bar{r} + B_m\tilde{q}_m(t). \quad (7)$$

By defining

$$\eta(t+1) = \tilde{\rho}(t+1) - \tilde{\rho}(t) \quad (8)$$

and

$$u(t+1) = r(t+1) - r(t), \quad (9)$$

and considering that in the steady state, $q_m(t)$ varies slowly, we obtain

$$\begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}(t+1) = \begin{bmatrix} I & A \\ 0 & A \end{bmatrix} \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}(t) + \begin{bmatrix} B_r \\ B_r \end{bmatrix} u(t), \quad (10)$$

which can be written in a more compact form as follows:

$$\underline{\eta}(t+1) = \underline{A}\underline{\eta}(t) + \underline{B}u(t). \quad (11)$$

We would like to synthesize a control law

$$u(t) = -K(t) \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}(t) \quad (12)$$

that minimizes the cost function

$$J = \frac{1}{2} \sum_{t=0}^{\infty} \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}^T(t) \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}(t) + u^T(t)Ru(t) \quad (13)$$

$$= \frac{1}{2} \sum_{t=0}^{\infty} \underline{\eta}^T(t) \underline{Q}\underline{\eta}(t) + u^T(t)Ru(t). \quad (14)$$

This problem can be easily solved using the algebraic Riccati equation. This approach is often referred to as the LQI (LQ with Integral action) method. It was used by Papageorgiou et al. (11) in designing a coordinated ramp-metering control algorithm for a freeway segment with 5 on-ramps.

The real control input (the metering rate) is given by

$$r(t) = r(t-1) - K(t) \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix}(t). \quad (15)$$

Due to the geometric constraints on the cell lengths, we have chosen the sampling time of the model to be 2 seconds. However, the control variable $r(t)$ can only be updated every 30 seconds for our I-210W test site, which is described in detail in Section 7. As a consequence, we have to employ a multirate approach to the LQI method.

We assume that the actual control input $r(t)$ is updated every p shortest time periods (STPs). $\text{STP} = 2$ s and $p = 15$ in our problem. The virtual control $u(t)$ as defined in (9) can only be

injected in the expanded system (11) every p STPs, i.e., $u(t)$ is nonzero only once every p STPs. Equivalently, instead of letting $u(t)$ be zero, we can let $\underline{B}(t)$ be zero and have a periodic dynamic system (11) in which

$$\underline{B}(t) = \begin{cases} \underline{B}, & \text{when } t = np \text{ for some } n \in \mathbb{Z}, \\ 0, & \text{when } t \neq np \text{ for any } n \in \mathbb{Z}. \end{cases} \quad (16)$$

In order to synthesize the optimal controller (12) for this periodic system, we have to solve the periodic algebraic Riccati equation

$$\begin{aligned} \underline{P}(t) = & \underline{A}^T(t)\underline{P}(t+1)\underline{A}(t) - \underline{A}^T(t)\underline{P}(t+1)\underline{B}(t)(R(t) \\ & + \underline{B}^T(t)\underline{P}(t+1)\underline{B}(t))^{-1} \underline{B}^T(t)\underline{P}(t+1)\underline{A}(t) + \underline{Q}(t) \end{aligned} \quad (17)$$

for a periodic solution

$$\underline{P}(t+p) = \underline{P}(t). \quad (18)$$

And the optimal control gain will be given by

$$K(t) = (R(t) + \underline{B}^T(t)\underline{P}(t+1)\underline{B}(t))^{-1} \underline{B}^T(t)\underline{P}(t+1)\underline{A}(t). \quad (19)$$

Note that $K(t)$ is periodic and is nonzero only at $t = np$. It can be computed off-line.

Another issue that needs to be addressed is saturation. The districts in the California State Department of Transportation (Caltrans) have established acceptable maximum and minimum metering rates. The ‘‘one vehicle per green per lane’’ policy enforced on the I-210W test site and typical driver and vehicle response times determine the green phase length of the metering signal to be 2 seconds. Similarly, the minimum length of the red phase is also 2 seconds. In addition, to avoid the increased probability of signal violation when drivers wait too long, the red phase length is usually limited to 18 or 13 seconds. Therefore, the maximum metering rate is usually 900 vehicles per hour per metered lane (vphpl), and the minimum metering rate is 180 or 240 vphpl.

On the other hand, the realized on-ramp entering flow may be less than the set metering rate, when the arrival rate of the vehicles (the demand) is less than the metering rate and there are not a sufficient number of vehicles waiting in the queue. Therefore, the control input $r(t)$ has to be saturated to not only the established maximum and minimum values, but also the number of available vehicles. In order to increase the response time of the integral control, some sort of ‘‘anti-windup’’ scheme has to be implemented.

Similarly to ALINEA (4), we distinguish the desired metering rate $r_c(t)$, which is set by the controller, from the realized on-ramp flow $r(t)$, which is measured by the entrance loop detector, as illustrated in Fig. 3, and modify the control algorithm (15) to be

$$r_c(t) = r(t-1) - K(t) \begin{bmatrix} \tilde{\rho} \\ \eta \end{bmatrix} (t), \quad (20)$$

and

$$r_c(t) = \min\{r_{\max}, \max\{r_{\min}, r_c(t)\}\}. \quad (21)$$

It has to be noted that using the actual ramp flow $r(t-1)$ in the integral control law (20) may lead to a non-zero offset in case of biased realization of the ramp flow due to, for example, signal violation (4).

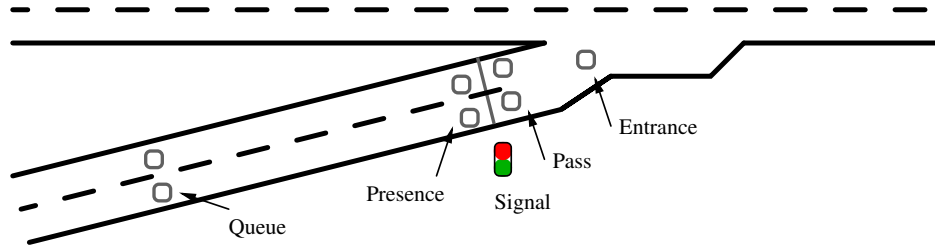


FIGURE 3 A typical configuration of loop detectors and signals on an on-ramp.

5 ON-RAMP QUEUE LENGTH REGULATION AND ESTIMATION

5.1 Queue Length Regulation

Whenever the on-ramp demand $d(t)$ exceeds the desired metering rate $r_c(t)$, a queue will form. However, the storage capacity of an on-ramp is often very limited. Without proper control, the vehicle queue length will quickly exceed this capacity, causing the vehicles to spill over into the surface streets and interfere with street traffic.

The current practice to regulate the queue length on California freeways is the so-called *queue-override*. A typical loop detector and signal configuration of an on-ramp is shown in Fig. 3. In the queue-override scheme, the signal controller compares the occupancy measured by the queue detectors with a threshold and determines whether the queue has reached the queue detectors. If the queue detector occupancy is above the threshold, the metering rate is increased by a certain level, e.g., 120 vphpl, every metering-rate update time interval (e.g., 30 seconds). When the queue detector occupancy falls below the threshold, the metering rate is reset to the value determined by the controller. This queue-override scheme can be viewed as an integral control with a saturated integrating rate and resetting. It has been noted (12–15) that this queue-override scheme leads to oscillatory behavior and under-utilization of on-ramp storage capacities. Gordon (12) attempted to improve the queue-override performance by filtering the occupancy signal and reducing the sampling time interval. Ozbay et al. (13, 14) showed through Paramics simulation that an integral ramp queue control can sometimes be effective to improve system performance. Smaragdis and Papageorgiou (15) proposed a proportional controller that relies on the on-ramp vehicle demands. Unfortunately, real-time on-ramp vehicle demand measurements are currently not available in the field, and historical demands may have to be used to realize the controller proposed in (15). The queue detector may provide an approximate demand measurement before the queue reaches it. After the queue extends beyond the queue detector, it will no longer be able to provide this measurement. The queue length regulator presented here does not require prior knowledge of the demands. It utilizes a queue length estimation algorithm to implement a proportional-integral (PI) controller.

The queue length dynamics is modeled as a simple integrator:

$$l(t+1) = l(t) + T_s(d(t) - r(t)) \quad \text{subject to } l \geq 0, \quad (22)$$

where $l(t)$ is the queue length (in number of vehicles), $d(t)$ is the demand (the flow entering the queue), and $r(t)$ is the on-ramp flow entering the mainline (the flow leaving the queue).

As previously discussed, the currently used queue-override scheme can be viewed as an integral control with a saturated integrating rate and resetting. A simple analysis shows that the integral control alone on an integrator is not asymptotically stable and a proportional action is

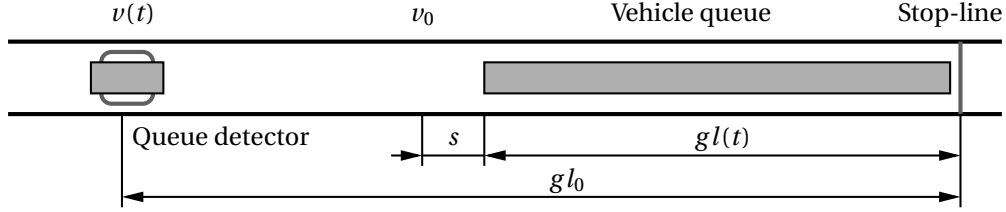


FIGURE 4 A schematics for on-ramp queue length estimation.

needed, i.e.,

$$r_c(z) = \left(k_P + \frac{k_I}{z-1} \right) \tilde{l}(z), \quad (23)$$

which is written in transfer function form. In (23), $\tilde{l}(z)$ is the queue length error. In the design of the PI queue length regulator, we treat $d(t)$ as an external disturbance and $r(t)$ as the control variable. The closed-loop sensitivity function from the disturbance to the error is

$$\frac{\tilde{l}(z)}{d(z)} = \frac{T_s(z-1)}{(z-1)^2 - k_P T_s z + (k_I - k_P) T_s}. \quad (24)$$

Proper PI gains k_P and k_I can be selected using the root locus method. Furthermore, the anti-windup and saturating mechanisms in (20) and (21) need to be implemented in the queue length regulator too.

5.2 Queue Length Estimation

Though it has a more stable response than the queue-override scheme, the PI regulator described in Section 5.1 needs the current queue length as its feedback, which unfortunately is not available in the field currently. A suitable estimator has to be designed using available information, such as the vehicle speed measured by the queue detector.

We assume the following simplified driving behavior model for a vehicle approaching the end of the queue: the vehicle decelerates at a constant rate, $-a$, from its cruising speed to a target speed v_0 at the position where the distance from the end of the queue is s . We also assume a uniform effective vehicle length g . Let l_0 be the number of vehicle spaces from the stop line to the queue detector and $v(t)$ be the vehicle speed measured by the queue detector. See Fig. 4.

A straightforward kinematic calculation yields

$$g(l_0 - l(t)) - s = \frac{v(t)^2 - v_0^2}{2a}, \quad (25)$$

where $l(t)$ is the current queue length, in number of vehicles. From (25), we obtain

$$gl(t) = gl_0 - s + \frac{v_0^2}{2a} - \frac{v(t)^2}{2a} = c_0 - c_2 v(t)^2. \quad (26)$$

To determine the coefficients c_0 and c_2 in (26), a curve fitting was performed on the $gl(t)$ and $v(t)$ data collected using the VISSIM (16) microscopic traffic simulator. Fig. 5 shows a typical scatter plot of queue lengths versus speeds. A few points need to be noted:

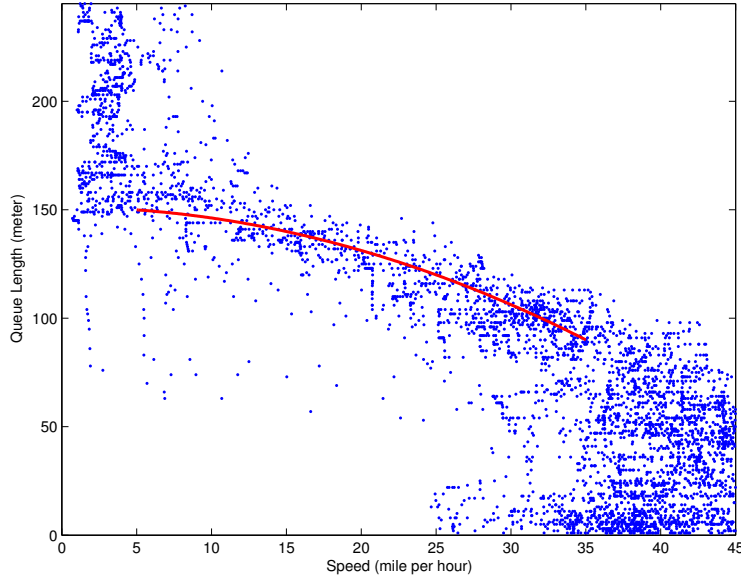


FIGURE 5 A scatter plot of queue lengths vs. queue detector speeds and the least median-of-squares curve fit for one of the on-ramps.

1. When the queue is shorter than a certain length, the approaching vehicles pass the queue detector at the drivers' desired cruising speeds, which are independent of the queue length. This phenomenon corresponds to the data points at the lower-right corner of the scatter plot.
2. When the queue is longer than gl_0 , i.e., the queue has extended beyond the queue detector, the measured speed is also a constant, which is related to the queue discharging rate and the vehicle lengths, and is also independent of the queue length. This phenomenon corresponds to the data points at the upper-left corner of the scatter plot.
3. There are many outliers among the data points. Therefore, the usual least-squares curve fitting method, which is biased toward outliers, is not suitable.

For these reasons, we neglected the data points whose speeds are below v_{\min} or above v_{\max} and those whose queue lengths are below l_{\min} or above l_{\max} in the curve fitting. These values were determined by visual inspection of the scatter plots.

To increase robustness to outliers, we used the least median-of-squares (17) curve fitting method, instead of the usual least (sum-of-)squares. The fitted curve is also shown in Fig. 5.

After the $l-v$ curve is fitted for each on-ramp, the difference between the actual and desired queue length, which is used as the feedback to the regulator (23), is estimated as

$$\tilde{l}(t) = \begin{cases} (c_0 - gl_0 - c_2v(t)^2)/g, & \text{if } v(t) \geq v_{\min}, \\ -kc_2(v(t)^2 - v_{\min}^2)/g, & \text{if } v(t) < v_{\min}, \end{cases} \quad (27)$$

where k is a tuning parameter.

When $v(t) < v_{\min}$, the end of the queue is very close to or beyond the queue detector, and the speed $v(t)$ measured by the queue detector is a constant, which is roughly gr_c . Therefore, (27) can be thought of as saturating \tilde{l} to $-kc_2((gr_c)^2 - v_{\min}^2)/g$, which is larger when the metering rate r_c is lower. This has a desirable effect on the regulator: The metering rate r_c will be increased more

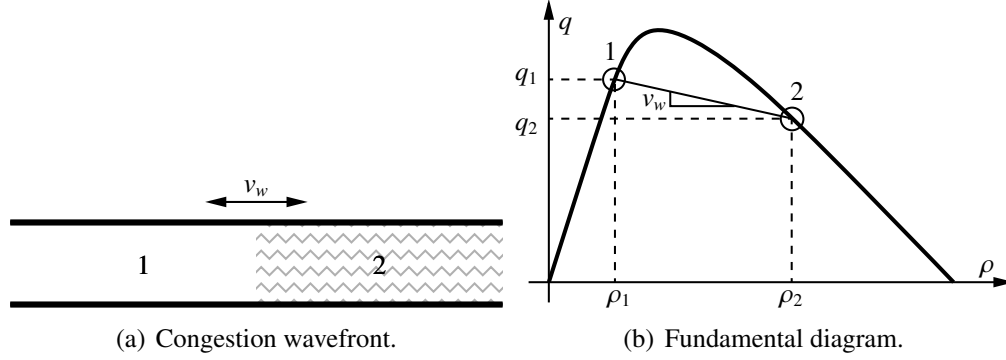


FIGURE 6 The fundamental diagram and congestion wavefront propagation.

aggressively when there is more room for this increase, and more slowly when r_c is close to its maximum value. In addition, this saturation value can be further tuned by changing the value of k .

It is also worth mentioning that the coefficients c_0 and c_2 identified by the least median-of-squares fitting are very close to the nominal values predicted by using the actual distance between the stop-line and the queue detector and a nominal vehicle deceleration of 2.5 m/s^2 . Therefore, when the queue length measurements are unavailable through any means to perform a curve fitting, these nominal values can be used in the queue length estimation.

6 A LOCALIZED RAMP-METERING STRATEGY

In addition to the switching LQI mainline traffic responsive metering controller and the queue regulator, a control strategy is still needed to achieve the best freeway performance. For example, the desired mainline density profiles and the desired queue lengths need to be defined. Gomes and Horowitz (18, 19) have formulated a linear program that minimizes the so-called *generalized total travel time* with explicit queue length constraints. This linear program is based on a modified version of the cell transmission model and produces a set of mainline densities, time-of-day ramp-metering rates and queue lengths as the solution, using the predicted traffic demands. These quantities can be used as the desired profiles in the switching controller and queue length regulator.

However, a localized ramp-metering strategy is also desired for other reasons including reduced algorithmic complexity, lower computational requirements, and higher robustness to changing traffic conditions such as unpredicted demands.

The main goal of freeway traffic control is to limit the spatial and temporal span of the congestion, i.e., to slow down congestion propagation in the upstream direction and to speed up any congestion wavefront that is moving downstream, in other words, to make the wavefront propagation speed v_w larger.

From the kinematic-wave traffic theory, for the congestion wavefront and freeway fundamental diagram (flow–density relationship) shown in Figures 6(a) and 6(b), the propagation speed of the the congestion wavefront is

$$v_w = -\frac{q_2 - q_1}{\rho_2 - \rho_1}, \quad (28)$$

where q_1 and q_2 are the flows at points 1 and 2, and ρ_1 and ρ_2 are the densities at points 1 and 2.

It can be seen from the foregoing analysis that the local ramp-metering control should move the congested point (point 2) to the left on the fundamental diagram, i.e., decrease the vehicle density. On the other hand, if there is no local congestion, vehicles can be allowed to enter the



FIGURE 7 A map of the test segment of Interstate 210 Westbound, from Vernon Avenue to Fair Oaks Avenue, in Pasadena, California (Composed using the U.S. Census Bureau TIGER/Line® (20) data).

mainline as fast as possible, so long as they do not induce congestion. Therefore, we choose the set-point of the feedback ramp-metering controller, in both the *free-flow* and the *congested* mode, to be the critical density ρ_c , i.e., the density at which congestion is about to form.

In order to utilize all of the available storage capacity on the ramp, it is natural to choose the set-point for the queue length regulator to be the maximum allowed queue length. A long on-ramp queue is also desired for the purpose of deterring short-distance travelers from using the freeway, thus making the freeway capacity available to long-distance travelers.

It has to be noted, however, that the objectives of the queue length regulator and the mainline traffic responsive metering controller conflict with each other: when the mainline traffic is light, the ramp-metering controller allows vehicles to enter the mainline as fast as possible, resulting in a short queue; at the same time, the queue length regulator is trying to accumulate vehicles so as to form a long queue. When the mainline traffic is heavy, the ramp-metering controller only allows vehicles to enter at a slow rate and the queue extends quickly beyond the limit; the queue length regulator will have to increase the metering rate to dissipate the queue until it reaches an acceptable length.

Considering these factors, we choose to set the metering rate at the *larger* of the two values determined by the mainline traffic responsive metering controller and the queue length regulator. Smaragdis and Papageorgiou proposed the same formula in (15).

7 TEST SITE

A segment of Interstate 210 Westbound (I-210W) in Pasadena, California, as shown in a map in Fig. 7, has been selected as the testbed for new ramp-metering algorithms. It is approximately 14 miles long, from Vernon Avenue (Mile Post 39.159) to Fair Oaks Avenue (Mile Post 25.4), with 20 metered on-ramps, 1 uncontrolled freeway-to-freeway connector (I-605), and 18 off-ramps. In our previous efforts, microscopic (21) and macroscopic (7) traffic simulation tools have been calibrated to this test site.

8 RESULTS

The switching LQI mainline-traffic responsive metering controller and the queue length regulator were implemented and interfaced with the VISSIM microscopic traffic simulation model that has been calibrated to the I-210W test segment by (21). The localized control strategy described in Section 6 was used. Fig. 8 shows the congestion patterns, as determined by the mixture Kalman

TABLE 1 Performance Measures for the I-210 Test Segment under Different Ramp-Metering Algorithms (Q/R: Queue Estimation and Regulation; U/S: Upstream; D/S: Downstream)

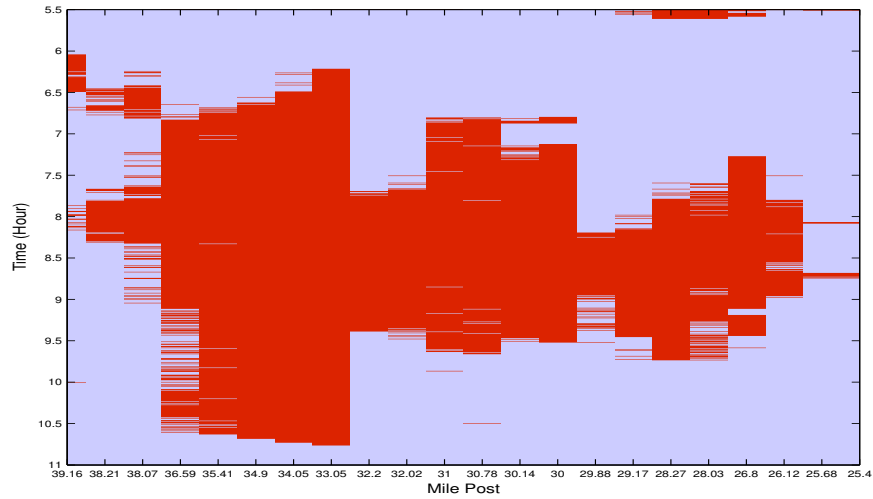
Mainline	Queue	TTD _V (10 ³ mile)	TTT _V (10 ³ hour)	TCD _V (10 ³ hour)	TTD _P (10 ⁶ mile)	TTT _P (10 ³ hour)	TCD _P (10 ³ hour)
None	None	973	24.0	8.52	1.32	31.7	10.8
Switching LQI	Q/R	972	22.6	7.19	1.32	29.9	9.0
Improvement		–	5.6%	16%	–	5.5%	16%
U/S ALINEA	Q/R	974	23.5	8.01	1.32	31.0	10.1
Improvement		–	2.1%	5.9%	–	2.1%	6.3%
D/S ALINEA	Q/R	973	23.3	7.82	1.32	30.7	9.8
Improvement		–	2.9%	8.2%	–	2.9%	8.6%
Switching LQI	None	974	22.3	6.81	1.32	29.3	8.3
Improvement		–	7.1%	20%	–	7.7%	23%
U/S ALINEA	None	973	22.3	6.87	1.32	29.3	8.39
Improvement		–	6.9%	19%	–	7.5%	22%
D/S ALINEA	None	974	22.5	7.04	1.32	29.5	8.59
Improvement		–	6.2%	17%	–	6.9%	20%

filtering (MKF) traffic state estimator (8, 9), before and after ramp metering. In the plots, red indicates *congested* mode and blue *free-flow*. The vertical axis is time, from 5:30 to 11:00 in the morning. The horizontal axis is the mile post along the freeway, and the traffic travels from left to right. It can be seen that the localized ramp metering strategy was able to reduce the congestion, in terms of both the spatial span and the time duration.

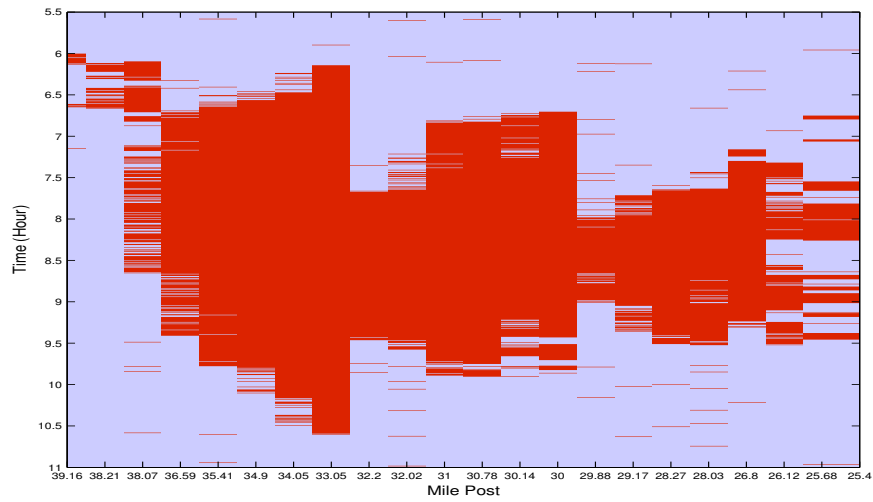
For these tests, the parameters in the multirate LQI design were set as follows: $Q = I$, $R = 5$ for the *free-flow* mode, and $Q = I$, $R = 20$ for the *congested* mode, while the gains $k_p = k_I = 120$ in the queue length regulator.

We also implemented the original and a modified version of ALINEA (4) and combined it with the queue length estimator and regulator that we have developed in Sections 5.2 and 5.1. In this modified ALINEA, we used the occupancy data measured upstream to the on-ramps, instead of those measured downstream in the original ALINEA, due to the loop-detector configuration on California freeways. (22) has shown, using the calibrated I-210W VISSIM model, that this modified ALINEA can achieve comparable, sometimes even better, performance, when compared to the original ALINEA. Optimal ALINEA gain (7000) and set-point (27% for upstream ALINEA and 21% for downstream ALINEA) found by (22) were used in our simulations. The same gain and set-point are used for all metered on-ramps.

Different ramp-metering algorithms, including 1) switching LQI, 2) the original upstream ALINEA, and 3) the modified downstream ALINEA, were tested with the I-210W VISSIM model. All the algorithms were tested both with and without the queue estimation and regulation algorithms. Under each scenario, 8 simulation runs were carried out, with 8 different VISSIM random seeds. The random seed was chosen to be the second of the computer clock at the time when it was changed, to ensure its randomness.



(a) Without ramp metering.



(b) With the proposed ramp metering algorithms.

FIGURE 8 Congestion modes for the I-210W test segment under different metering scenarios, blue: *free-flow*, and red: *congested*.

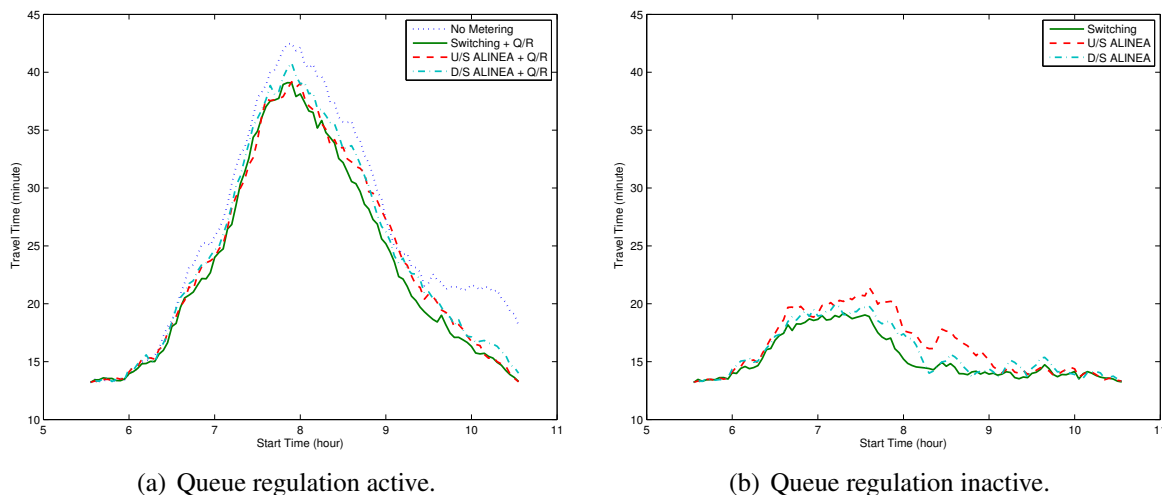


FIGURE 9 Individual mainline travel times as functions of the start time for different ramp-metering scenarios.

Some of the performance measures for this freeway segment, as defined in Section 2, are listed in Table 1. The listed numbers are the averages from the 8 simulation runs for each scenario. In calculating these quantities, the average passenger numbers per one low- and high-occupancy vehicle are assumed to be 1.2 and 2.5, respectively, and the nominal free-flow speed v_0 is 63 miles per hour.

Under all the scenarios, the freeway segment served almost the same amount of demand, as measured by the Total Vehicle Distance TTD_V or Total Passenger Distance TTD_P . Ramp-metering was able to reduce the congestion under all the metered scenarios. For example, with the switching LQI mainline control and queue length regulation, the Total Vehicle Delay (also known as Congestion Delay) TCD_V was reduced by 16%, while with the switching LQI mainline control only, TCD_V was reduced by 20%.

When only the switching mainline-traffic responsive metering was used, without activating the queue length regulator, on-ramp queues can be accumulated to arbitrary lengths, sometimes hundreds of vehicles. In this case, almost all the congestion on the mainline was eliminated. Another interesting phenomenon in this case is that the relative improvements in terms of passenger-weighted performance measures were greater than those in terms of vehicle performance measures. This is because many of the metered on-ramps on this freeway segment have designated lanes for HOVs to bypass the long queues.

It can also be seen from the numbers in Table 1 that the switching control algorithm outperforms both the original and the modified ALINEAs, whether the queue length estimator and regulator was active or not. The switching LQI controller distinguishes from the ALINEA algorithm in two aspects, which contribute to its better performance: 1) the switching LQI controller employs different feedback structures that are suitable to the different traffic dynamics under the *free-flow* or *congested* conditions; and 2) the proportional action in the LQI controller provides a faster reaction than the integral action alone.

Individual mainline travel times as functions of the start time for different ramp-metering scenarios are plotted in Fig. 9. It can be observed from the figure that under all the metered scenarios, the onset of the congestion was slower and the dissipation of the congestion was faster

than the un-metered scenario. In particular, the switching mainline control with queue regulation was able to dissipate congestion faster than the upstream and downstream ALINEAs. This is more obvious in Fig. 9(b), where the individual travel times are plotted for the scenarios when the queue length constraints were not enforced. The switching mainline control was able to achieve a lower maximum travel time and it dissipated the congestion much faster than the downstream and upstream ALINEAs did.

9 CONCLUSIONS

In this paper, we presented a novel switching traffic-responsive ramp-metering controller that adapts to the different traffic dynamics under free-flow or congested conditions. The approach of multirate LQI was used to compensate for disturbances and accommodate the difference in the model sampling time and the metering rate update interval. In addition, a PI queue length regulator was designed to prevent the on-ramp queue from exceeding the storage capacity and yield improved performance over the currently used *ad hoc* “queue-override” scheme. A queue length estimator was also designed to provide feedback to the queue length regulator, using the queue-detector speed data that are available in the field. A localized strategy was proposed to achieve the control goal of reducing the spatial and temporal extent of the congestion, using locally available information.

Test results on the calibrated VISSIM I-210W microscopic model demonstrated the performance and effectiveness of the switching ramp-metering controller, the queue length estimator and regulator, and the overall control strategy. The Total Vehicle and Passenger Delays were both reduced by 16%, while the Total Vehicle Time was improved by 5.6%. As a comparison, simulation results of ALINEA were also presented. The switching mainline-traffic responsive control was able to outperform ALINEA, when both algorithms were combined with the same queue length estimator and regulator.

ACKNOWLEDGMENT

This work was supported by California Partners for Advanced Transit and Highways (PATH) under Task Orders 4136 and 5503.

REFERENCES

1. Schrank, D. and T. Lomax. *The 2004 Urban Mobility Report*. Tech. rep., Texas Transportation Institute, Sep. 2004. URL <http://mobility.tamu.edu/>.
2. Cambridge Systematics, Inc. and Texas Transportation Institute. *Traffic Congestion and Reliability: Linking Solutions to Problems*. Tech. rep., Federal Highway Administration, U.S. Department of Transportation, Jul. 19, 2004. URL http://www.ops.fhwa.dot.gov/congestion_report/.
3. Muñoz, L., X. Sun, R. Horowitz, and L. Alvarez. Traffic Density Estimation with the Cell Transmission Model. In *Proceedings of the 2003 American Control Conference*. Denver, Colorado, USA, Jun. 2003, pp. 3750–3755.
4. Papageorgiou, M., H. Hadj-Salem, and J.-M. Blosseville. ALINEA: A Local Feedback Control Law for On-Ramp Metering. *Transportation Research Record*, , no. 1320, 1991, pp. 58–64.
5. Daganzo, C. F. The Cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamic Theory. *Transportation Research Part B: Methodological*, vol. 28, no. 4, Aug. 1994, pp. 269–287.

6. Daganzo, C. F. The Cell Transmission Model, Part II: Network Traffic. *Transportation Research Part B: Methodological*, vol. 29, no. 2, Apr. 1995, pp. 79–93.
7. Muñoz, L., X. Sun, D. Sun, G. Gomes, and R. Horowitz. Methodological Calibration of the Cell Transmission Model. In *Proceedings of the 2004 American Control Conference*. Boston, Massachusetts, USA, Jun. 30–Jul. 2, 2004, pp. 798–803.
8. Sun, X., L. Muñoz, and R. Horowitz. Highway Traffic State Estimation Using Improved Mixture Kalman Filters for Effective Ramp Metering Control. In *Proceedings of the 42nd IEEE Conference on Decision and Control*. Maui, Hawaii, USA, Dec. 9–12, 2003, pp. 6333–6338.
9. Sun, X., L. Muñoz, and R. Horowitz. Mixture Kalman Filter Based Highway Congestion Mode and Vehicle Density Estimator and its Application. In *Proceedings of the 2004 American Control Conference*. Boston, Massachusetts, USA, Jun. 30–Jul. 2, 2004, pp. 2098–2103.
10. Sun, X. and R. Horowitz. A Localized Switching Ramp-Metering Controller with a Queue Length Regulator for Congested Freeways. In *Proceedings of the 2005 American Control Conference*. Portland, Oregon, USA, Jun. 8–10, 2005, pp. 2141–2146.
11. Papageorgiou, M., J.-M. Blosseville, and 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 Part A: General*, vol. 24, no. 5, Sep. 1990, pp. 361–370.
12. Gordon, R. L. Algorithm for Controlling Spillback from Ramp Meters. *Transportation Research Record*, , no. 1554, 1996, pp. 162–171.
13. Ozbay, K. and P. Kachroo. *Feedback Ramp Metering for Intelligent Transportation System*. Kluwer Academics, New York, 2003.
14. Ozbay, K., I. Yasar, and P. Kachroo. Comprehensive Evaluation of Feedback-Based Freeway Ramp-Metering Strategy by Using Microscopic Simulation: Taking Ramp Queues into Account. *Transportation Research Record*, , no. 1867, 2004, pp. 89–96.
15. Smaragdis, E. and M. Papageorgiou. Series of New Local Ramp Metering Strategies. *Transportation Research Record*, , no. 1856, 2004, pp. 74–86.
16. PTV AG. VISSIM. Web page, 2004. URL http://www.english.ptv.de/cgi-bin/traffic/traf_vissim.pl.
17. Rousseeuw, P. J. Least Median of Squares Regression. *Journal of the American Statistical Association*, vol. 79, no. 388, 1984, pp. 871–880.
18. Gomes, G. and R. Horowitz. Globally Optimal Solutions to the Onramp Metering Problem, Part I. In *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems*. Washington, D.C., USA, Oct. 3–6, 2004, pp. 509–514.
19. Gomes, G. and R. Horowitz. Globally Optimal Solutions to the Onramp Metering Problem, Part II. In *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems*. Washington, D.C., USA, Oct. 3–6, 2004, pp. 515–520.
20. U.S. Census Bureau. TIGER®: Topologically Integrated Geographic Encoding and Referencing system. Web site, Jan. 2005. URL <http://www.census.gov/geo/www/tiger/>. 2004 First Edition.
21. Gomes, G., A. D. May, and R. Horowitz. Calibration of VISSIM for a Congested Highway. In *The 83rd Annual Meeting of the Transportation Research Board*. Washington, D.C., USA, Jan. 2004.
22. Gomes, G. C. *Optimization and Microsimulation of On-ramp Metering for Congested Freeways*. Ph.D. dissertation, University of California, Berkeley, 2004.