2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

A Localized Switching Ramp-Metering Controller with a Queue Length Regulator for Congested Freeways

Xiaotian Sun and Roberto Horowitz

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. 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 cell transmission model (CTM) based macroscopic traffic simulator have demonstrated the performance and effectiveness of the switching ramp-metering controller, the queue length regulator, and the control strategy. The switching metering controller tracks the desired mainline density profiles in both free-flow and congested conditions. Under the localized control strategy, the Total Vehicle Delay is reduced by 7.7%, while the Total Vehicle Time is reduced by 5.7%, in the CTM traffic simulator.**

I. 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 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 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 these controllers on a calibrated macroscopic traffic simulator are presented.

II. 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.

- $D_{\mathrm{V,tot}}$ Total Vehicle Distance, which is defined as the sum of the distances traveled by all the vehicles in L within T.
- $T_{\mathrm{V,tot}}$ Total Vehicle Time, which is the sum of the time that is spent by all vehicles in L within T. It includes the time spent while vehicles are waiting in the on-ramp queues.
- $DL_{\mathrm{V,tot}}$ Total Vehicle Delay, which is the difference between the Total Vehicle Time and the time that would be spent by all the vehicles if there were no congestion. $DL_{\mathrm{V,tot}} = T_{\mathrm{V,tot}} - D_{\mathrm{V,tot}}/v_0$, where v_0 is the nominal free-flow speed.
- $\bar{v}_{\mathrm{V,tot}}$ Average Total Vehicle Speed $\bar{v}_{\mathrm{V,tot}} = D_{\mathrm{V,tot}}/T_{\mathrm{V,tot}}$.

III. Switching-Mode Model and Controllability

In our previous work [3], we piecewise linearized the cell transmission model [4], [5] and derived a switching-mode traffic model. Depending on the traffic conditions in a

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

The authors are with the Department of Mechanical Engineering, University of California at Berkeley, Berkeley, CA 94720-1740. E-mail addresses: {sunx,horowitz}@me.berkeley.edu.

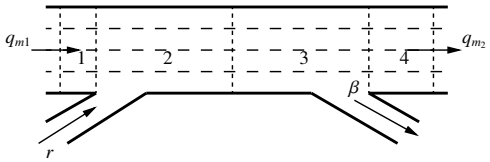


Fig. 1. A schematic plot of a 4-cell freeway section with one on-ramp and one off-ramp.

freeway section, such as the one shown in Fig. 1, the traffic is in a different mode, e.g., *free-flow* or *congested*.[1] 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,

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

$$= A(1)\rho(t) + B_q(1)q(t), \tag{2}$$

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,

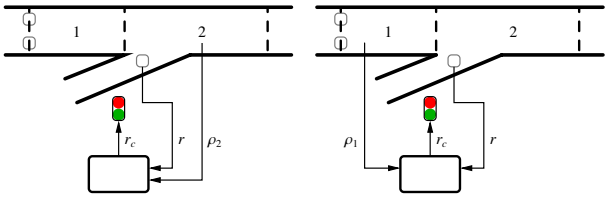
$$\begin{bmatrix}\rho_1\\ \rho_2\\ \rho_3\\ \rho_4\end{bmatrix}(t+1) = \begin{bmatrix} 1-\frac{w_{c1}T_s}{l_1} & \frac{w_{c2}T_s}{l_1} & 0 & 0 \\ 0 & 1-\frac{w_{c2}T_s}{l_2} & \frac{w_{c3}T_s}{l_2} & 0 \\ 0 & 0 & 1-\frac{w_{c3}T_s}{l_3} & \frac{1}{1-\beta}\frac{w_{c4}T_s}{l_3} \\ 0 & 0 & 0 & 1-\frac{w_{c4}T_s}{l_4} \end{bmatrix}\begin{bmatrix}\rho_1\\ \rho_2\\ \rho_3\\ \rho_4\end{bmatrix}(t) + \begin{bmatrix} 0 & 0 & \frac{T_s}{l_1} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -\frac{T_s}{l_4} & 0 \end{bmatrix}\begin{bmatrix} q_{m1}\\ q_{m2}\\ r\end{bmatrix}(t) + \begin{bmatrix} \frac{w_{c1}T_s}{l_1} & -\frac{w_{c2}T_s}{l_1} & 0 & 0 \\ 0 & \frac{w_{c2}T_s}{l_2} & -\frac{w_{c3}T_s}{l_2} & 0 \\ 0 & 0 & \frac{w_{c3}T_s}{l_3} & -\frac{1}{1-\beta}\frac{w_{c4}T_s}{l_3} \\ 0 & 0 & 0 & \frac{w_{c4}T_s}{l_4} \end{bmatrix}\begin{bmatrix}\rho_{J1}\\ \rho_{J2}\\ \rho_{J3}\\ \rho_{J4}\end{bmatrix}, \tag{3}$$

$$= A(2)\rho(t) + B_q(2)q(t) + B_J(2)\rho_J, \tag{4}$$

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.

[1]This two-mode (*free-flow* and *congested*) setup is an approximation. In reality, it is possible to have mixed cases, such as the mode wherein half of the cells in a section are in free-flow and half are in congestion. These mixed cases are neglected here, because 1) the divided freeway sections are short and the mixed cases are often transient, and 2) the traffic state estimator [6], [7] has provided satisfactory results using only these two possible modes.



(a) *Free-flow* mode. (b) *Congested* mode.

Fig. 2. Different control structures for different congestion modes.

This observation is critical to the design of an on-ramp metering algorithm, because it determines a 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 [6], [7] that is able to accurately estimate these quantities in real time. The estimated congestion mode is used to determine the appropriate control structure, and the estimated vehicle densities are used as feedback.

IV. Multirate LQI Design for the Mainline Traffic Responsive Ramp-Metering Controller

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

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

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

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

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

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

Let

$$z(t+1) = \tilde{\rho}(t+1) - \tilde{\rho}(t), \tag{8}$$

and

$$u(t+1) = r(t+1) - r(t). \tag{9}$$

We then have

$$z(t+1) = Az(t) + B_r u(t) + B_m(\tilde{q}_m(t) - \tilde{q}_m(t-1)), \quad (10)$$

and

$$\tilde{\rho}(t+1) = \rho(t) + Az(t) + B_r u(t) + B_m(\tilde{q}_m(t) - \tilde{q}_m(t-1)). \quad (11)$$

Considering that in the steady state, $q_m(t)$ varies slowly, we neglect the $B_m(\tilde{q}_m(t) - \tilde{q}_m(t-1))$ terms and combine the two equations together to obtain

$$\begin{bmatrix} \tilde{\rho} \\ z \end{bmatrix}(t+1) = \begin{bmatrix} I & A \\ 0 & A \end{bmatrix}\begin{bmatrix} \tilde{\rho} \\ z \end{bmatrix}(t) + \begin{bmatrix} B_r \\ B_r \end{bmatrix} u(t), \quad (12)$$

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

$$\underline{z}(t+1) = \underline{A}\underline{z}(t) + \underline{B}u(t). \quad (13)$$

We would like to synthesize a control law

$$u(t) = -K(t)\begin{bmatrix} \tilde{\rho} \\ z \end{bmatrix}(t) \quad (14)$$

that minimizes the cost function

$$J = \frac{1}{2}\sum_{t=0}^{\infty}\begin{bmatrix} \tilde{\rho} \\ z \end{bmatrix}^H(t)\begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \tilde{\rho} \\ z \end{bmatrix}(t) + u^H(t)Ru(t) \quad (15)$$

$$= \frac{1}{2}\sum_{t=0}^{\infty}\underline{z}^H(t)\underline{Q}\underline{z}(t) + u^H(t)Ru(t). \quad (16)$$

This problem can be easily solved using the algebraic Riccatti equation. This approach is often referred to as the LQI (LQ with Integral action) method. It was used by Papageorgiou et al. [8] 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

$$r(t) = r(t-1) - K(t)\begin{bmatrix} \tilde{\rho} \\ z \end{bmatrix}(t). \quad (17)$$

Due to the geometric constraints on the cell lengths, we have chosen the sampling time of the model to be 2 seconds. However, in the field, the control variable $r(t)$ can only be updated every 30 seconds. 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). 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 (13) 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 (13) in which

$$\underline{B}(t) = \begin{cases} \underline{B}, & \text{when } t = np \text{ for some } n \in \mathbb{Z}, \\ 0, & \text{when } t \neq np \text{ for any } n \in \mathbb{Z}. \end{cases} \quad (18)$$

In order to synthesize the optimal controller (14) for this periodic system, we have to solve the periodic algebraic Riccati equation

$$\underline{P}(t) = \underline{A}^H(t)\underline{P}(t+1)\underline{A}(t) - \underline{A}^H(t)\underline{P}(t+1)\underline{B}(t)\Big(R(t) + \underline{B}^H(t)\underline{P}(t+1)\underline{B}(t)\Big)^{-1}\underline{B}^H(t)\underline{P}(t+1)\underline{A}(t) + \underline{Q}(t) \quad (19)$$

for a periodic solution

$$\underline{P}(t+p) = \underline{P}(t). \quad (20)$$

And the optimal control gain will be given by

$$K(t) = (R(t) + \underline{B}^H(t)\underline{P}(t+1)\underline{B}(t))^{-1}\underline{B}^H(t)\underline{P}(t+1)\underline{A}(t). \quad (21)$$

Given $\underline{P}(p)$, it is straightforward to calculate

$$\underline{P}(1) = (I + \underline{A} + \cdots + \underline{A}^{p-2})^H\underline{Q}(I + \underline{A} + \cdots + \underline{A}^{p-2}) + (\underline{A}^{p-1})^H\underline{P}(p)(\underline{A}^{p-1}). \quad (22)$$

Then we have

$$\underline{P}(p) = \underline{Q}_p + \underline{A}_p^H\underline{P}(p)\underline{A}_p - (\underline{A}_p^H\underline{P}(p)\underline{B}_p + S_p)\Big(R_p + \underline{B}_p^H\underline{P}(p)\underline{B}_p\Big)^{-1}(\underline{A}_p^H\underline{P}(p)\underline{B}_p + S_p)^H, \quad (23)$$

where

$$\underline{A}_p = \underline{A}^p,$$
$$\underline{B}_p = \underline{A}^{p-1}\underline{B},$$
$$\underline{Q}_p = \Big(\sum_{i=0}^{p-1}A^i\Big)^H\underline{Q}\Big(\sum_{i=0}^{p-1}A^i\Big),$$
$$R_p = R + \underline{B}^H\Big(\sum_{i=0}^{p-2}A^i\Big)^H\underline{Q}\Big(\sum_{i=0}^{p-2}A^i\Big)\underline{B},$$

and

$$S_p = \underline{A}^H\Big(\sum_{i=0}^{p-2}A^i\Big)^H\underline{Q}\Big(\sum_{i=0}^{p-2}A^i\Big)\underline{B}.$$

Note that (23) is in the standard form of a discrete-time algebraic Riccati equation. The resulting periodic control gain is given by

$$K(t) = \begin{cases} K_p, & \text{when } t = np \text{ for some } n \in \mathbb{Z}, \\ 0, & \text{when } t \neq np \text{ for any } n \in \mathbb{Z}, \end{cases} \quad (24)$$

where

$$K_p = \Big(R_p + \underline{B}_p^H\underline{P}(p)\underline{B}_p\Big)^{-1}\Big(\underline{B}_p^H\underline{P}(p)\underline{A}_p + S_p\Big).$$

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" policy 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

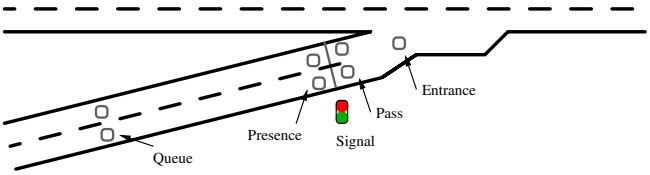


Fig. 3. A typical configuration of loop detectors and signals on an on-ramp.

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 [9], 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 (17) to be

$$r_c(t) = r(t-1) - K(t)\begin{bmatrix}\tilde{\rho}\\ z\end{bmatrix}(t), \tag{25}$$

and

$$r_c(t) = \min\{r_{\max}, \max\{r_{\min}, r_c(t)\}\}. \tag{26}$$

V. An On-Ramp Queue Length Regulator

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 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 [10], [11] that this queue-override scheme leads to oscillatory behavior and under-utilization of on-ramp storage capacities. Gordon [10] attempted to improve the queue-override performance by filtering the occupancy signal and reducing the sampling time interval. Smaragdis and Papageorgiou [11] 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 [11].

The queue length dynamics is modeled as a simple integrator:

$$l(t+1) = l(t) + T_s(d(t) - r(t)) \quad \text{subject to } l \geq 0, \tag{27}$$

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 needed, i.e.,

$$r_c(z) = \left(k_P + \frac{k_I}{z-1}\right)\tilde{l}(z), \tag{28}$$

which is written in transfer function form. In (28), $\tilde{l}(z)$ 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

$$\frac{\tilde{l}(z)}{d(z)} = \frac{T_s(z-1)}{(z-1)^2 - k_P T_s z + (k_I - k_P)T_s}. \tag{29}$$

Proper PI gains k_P and k_I can be selected using the root locus method. Furthermore, the anti-windup and saturating mechanisms in (25) and (26) need to be implemented in the queue length regulator too.

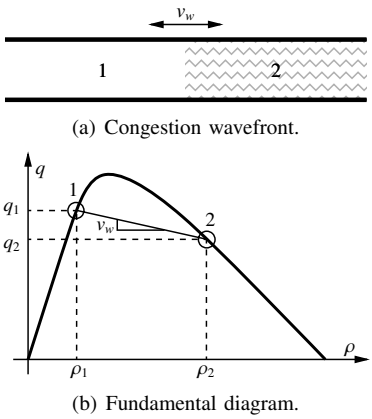
In the field, the queue length feedback that is needed by the regulator is not generally available. However, in this paper, we assume that we have this measurement, since the macroscopic traffic simulator is able to provide it. We have designed an estimator for the queue length, using the available field quantities [12].

VI. A Localized Ramp-Metering Strategy

In addition to the switching 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 [13], [14] 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



(a) Congestion wavefront.

(b) Fundamental diagram.

Fig. 4. The fundamental diagram and congestion wavefront propagation.

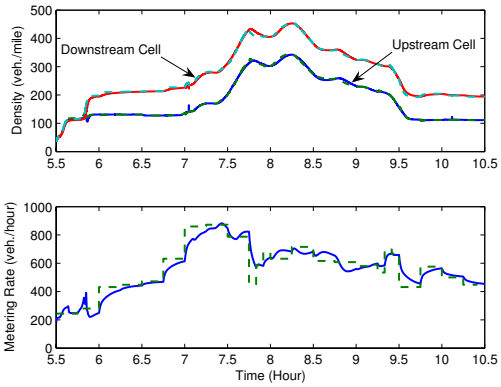


Fig. 5. Time varying desired mainline density tracking with the switching ramp-metering controller. Simulated with the calibrated cell transmission model.

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 kinetic-wave traffic theory, for the congestion wavefront and freeway fundamental diagram (flow–density relationship) shown in Figures 4(a) and 4(b), the propagation speed of the the congestion wavefront is

$$v_w = -\frac{q_2 - q_1}{\rho_2 - \rho_1}, \tag{30}$$

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 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 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 [11].

VII. Test Site

A segment of Interstate 210 Westbound (I-210W) in Pasadena, California, 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 [15] and macroscopic [16] traffic simulation tools have been calibrated to this test site.

VIII. Results

The local mainline traffic responsive ramp-metering controller, without the queue length regulator, was first tested on the calibrated macroscopic cell transmission model (CTM) [16]. The desired mainline vehicle densities were obtained from an open-loop simulation of the CTM with a set of optimal time-of-day metering rates solved by the linear program in [13], [14]. Fig. 5 shows an example plot of the tracking performance of the mainline vehicle densities in the upstream and downstream cells (1 and 2 in Fig. 2) and the convergence of the ramp-metering rate. In the plot, the dashed lines are the desired densities or metering rate, and the solid lines are the actual quantities.

The mainline traffic responsive metering controller and the queue length regulator, together with the localized control strategy described in Section VI, were also tested in the calibrated macroscopic CTM simulator [16]. In addition to the regular traffic quantities that are measured by the loop detectors, this CTM simulator also provides the actual on-ramp queue lengths, which were used in the queue length regulator as the feedback. As a comparison, the optimal time-of-day metering plan, as determined by the algorithms proposed in [13], [14], was also simulated with the CTM simulator.

TABLE I

PERFORMANCE MEASURES WITH DIFFERENT RAMP METERING STRATEGIES ON THE CALIBRATED I-210W CELL TRANSMISSION MODEL.

	$D_{\text{V,tot}}$ (10^3 mile)	$T_{\text{V,tot}}$ (10^3 hour)	$DL_{\text{V,tot}}$ (10^3 hour)	$\bar{v}_{\text{V,tot}}$ (mph)
No metering	535.9	31.46	22.95	17.0
Localized strategy	535.5	29.68	21.18	18.0
Improvement	–	5.7%	7.7%	5.9%
Optimal plan	535.2	29.86	21.36	17.9
Improvement	–	5.1%	6.9%	5.3%

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.

Some freeway performance measures in these simulations, as defined in Section II, are listed in Table I. In calculating the delays, the nominal free-flow speed v_0 is assumed to be 63 miles per hour.

In all simulations, the freeway segment served almost the same amount of demand, as measured by the Total Vehicle Distance $D_{\text{V,tot}}$. However, traffic-responsive ramp metering with the localized strategy was able to reduce the congestion on the mainline and increase the Average Total Vehicle Speed $\bar{v}_{\text{V,tot}}$ by almost 6%, while the optimal metering plan increased $\bar{v}_{\text{V,tot}}$ by 5.3%. The Total Vehicle Delay $DL_{\text{V,tot}}$, which is the extra time spent due to the congestion, was reduced by 7.7% and 6.9% by the localized strategy and by the optimal metering plan, respectively, while the Total Vehicle Time $T_{\text{V,tot}}$ was reduced by 5.7% and 5.1% by these two metering strategies, respectively.

From these numbers, it can be seen that the localized metering strategy had even slightly better performance than the optimal metering plan. This is partly due to the 5-minute time resolution of the optimal plan, which was chosen for reasons of computational efficiency [13], [14].

It should be pointed out that the purpose of the macroscopic simulations is not to extensively evaluate the performance of the proposed ramp-metering algorithms on a real freeway. Instead, these macroscopic simulations provide a means to examine how well the ramp-metering algorithms that were designed based on the piecewise linearized switching-mode model perform on the full CTM. For microscopic simulation results with a calibrated VISSIM model, please see [12].

IX. 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 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 macroscopic traffic simulator demonstrated the performance and effectiveness of the switching mainline traffic responsive ramp-metering controller, the queue length regulator and the control strategy. The Total Vehicle Delay was reduced 7.7%, while the Total Vehicle Time was reduced by 5.7%.

REFERENCES

[1] D. Schrank and T. Lomax, "The 2004 urban mobility report," Texas Transportation Institute, Tech. Rep., Sept. 2004. [Online]. Available: http://mobility.tamu.edu/

[2] Cambridge Systematics, Inc. and Texas Transportation Institute, "Traffic congestion and reliability: Linking solutions to problems," Federal Highway Administration, Tech. Rep., July 19, 2004. [Online]. Available: http://www.ops.fhwa.dot.gov/congestion_report/

[3] L. Muñoz, X. Sun, R. Horowitz, and L. Alvarez, "Traffic density estimation with the cell transmission model," in *Proceedings of the 2003 American Control Conference*, Denver, CO, USA, June 2003, pp. 3750–3755.

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

[5] ——, "The cell transmission model, part II: Network traffic," *Transportation Research Part B: Methodological*, vol. 29, no. 2, pp. 79–93, Apr. 1995.

[6] X. Sun, 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.

[7] ——, "Mixture Kalman filter based highway congestion mode and vehicle density estimator and its application," in *Proceedings of the 2004 American Control Conference*, Boston, MA, USA, June 30–July 2, 2004, pp. 2098–2103.

[8] M. Papageorgiou, 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, pp. 361–370, Sept. 1990.

[9] M. Papageorgiou, H. Hadj-Salem, and J.-M. Blosseville, "ALINEA: A local feedback control law for on-ramp metering," *Transportation Research Record*, no. 1320, pp. 58–64, 1991.

[10] R. L. Gordon, "Algorithm for controlling spillback from ramp meters," *Transportation Research Record*, no. 1554, pp. 162–171, 1996.

[11] E. Smaragdis and M. Papageorgiou, "Series of new local ramp metering strategies," *Transportation Research Record*, no. 1856, pp. 74–86, 2004.

[12] X. Sun and R. Horowitz, "Localized switching ramp-metering control with queue length estimation and regulation and microscopic simulation results," in *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, July 4–8, 2005, to appear.

[13] G. Gomes 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., Oct. 3–6, 2004, pp. 509–514.

[14] ——, "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., Oct. 3–6, 2004, pp. 515–520.

[15] G. Gomes, 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., Jan. 2004.

[16] L. Muñoz, 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, MA, USA, June 30–July 2, 2004, pp. 798–803.