IMPUTATION OF RAMP FLOW DATA FOR FREEWAY TRAFFIC SIMULATION

Ajith Muralidharan
Department of Mechanical Engineering
University of California,
Berkeley CA 94720
Phone: (510) 642-5109
Email: ajith@berkeley.edu.
(Corresponding author)

Roberto Horowitz
Department of Mechanical Engineering
University of California,
Berkeley CA 94720
Phone: (510) 642-4675, Fax:(510) 643-5599
Email: horowitz@berkeley.edu.

Length = 3000 Words + 6 Figures + 1 Table = 4750 words

August 1, 2008

Abstract

The Tools for Operations Planning (TOPL) employs the Link-Node Cell Transmission model for macroscopic freeway traffic simulations for specifying operational strategies like ramp metering, demand and incident management. Traffic flow and occupancy data from loop detectors is used for calibrating these models and specifying the inputs to the simulation. However, flow data from ramps are often to be found missing or incorrect. This paper elaborates an imputation procedure used to determine these ramp flows. This automated imputation procedure is based on an adaptive identification technique that tries to minimize the error between the simulated and the measured densities. The simulation results using the imputed flow data indicate good conformation with loop detector measurements.

1 INTRODUCTION

Tools for operational planning (TOPL) is a suite of tools used for (1) specifying operational improvements - ramp metering, incident management, traveler information and demand management and (2) quickly estimating the benefits such improvements are likely to provide. This is an essential component of the California Department of Transportation (Caltrans) "corridor management program" - which was introduced to reduce the congestion in 2025 by 40 percent [1].
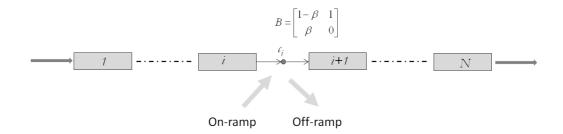
TOPL is based on the macroscopic Link-Node Cell Transmission Model (CTM) [2]. Traditionally, transportation planning investigations favor use of microscopic models. However, data collection and model calibration efforts are significant for microscopic models, therby slowing these efforts [3].The CTM model is based on aggregate variables such as volume (flow) and density. This data is available for California Freeways from vehicle detector stations (vds), which contain loop detectors. PeMS [4] routinely archives the flow, occupancy and speed data from these vds. In general, the CTM requires the flow and density measurements from the mainline vds (positioned along the freeway) for calibration of the fundamental diagrams. The ramp flows need to be specified as an input for the simulations. However, the data from the ramps are often found to be missing or incorrect. Hence, it becomes essential to impute the onramp and offramp flows to completely specify the simulation model.

This paper illustrates an imputation procudure for determining ramp flows using the Link Node CTM . The ramp flows are determined using an adaptive identification technique which tries to minimize the error between the model calculated densities and the measurements. Section 2 reviews the Link-Node Cell Transmission model used for traffic flow simulations. It also derives a simple four-state switching model used for freeway-corridor simulations. Section 3 explains the imputation procedure used for determining onramp and offramp flows. Finally, section 4 illustrates an example where the imputation procedure is used to specify inputs for a 26-mile long I-210E freeway in the Los Angeles area.

2   LINK NODE CELL TRANSMISSION MODEL

The Link-Node Cell transmission model (LN-CTM) is used to simulate traffic flows in traffic networks. Aurora, a simulation tool in TOPL, is based on this CTM implementation [5]. Other implementations of the CTM include the Assymetric Cell Transmission Model (ACTM) used particulary for freeway traffic simulation [6]. The Link-Node Cell transmission model is preferred for simulation, since it has the capability to simulate traffic networks which include freeways and arterial networks, as compared to the ACTM, which has been primarily used for freeway simulations. As a result, the Link-Node CTM model has been used for imputation of on-ramp and off-ramp flows in this paper.

In the LN-CTM model, the freeway (or any traffic network) is specified by a graph of links. Links represent a road segment, which carries traffic. Nodes are formed at the junction of Links, where traffic flow exchange takes place. The flow exchange is indicated by a time varying split-ratio matrix, which specifies the portion of traffic moving from a particular input link to an output link. While a normal link connects two Nodes, a "source" link is used to introduce traffic into the network whereas a "sink" is used to accept traffic moving out of the network . A source link implements a queue model.

$$B = \begin{bmatrix} 1-\beta & 1 \\ \beta & 0 \end{bmatrix}$$

FIGURE. 1 —— Freeway with N links. Each Node contains a maximum of one on- and one off-ramp

Figure 1 shows the freeway divided specified in the Link-Node framework. Each Node contains a maximum of one on- and one off-ramp. Freeflow prevails in both the boundaries of the freeway. Vehicles enter through a "source" attached to the upstream cell. The onramps are also represented as source links, while the offramps are represented as sinks. It is also assumed that the off-ramps are in freeflow, i.e., the flow to the off-ramps are not restricted by their flow capacity or space restrictions. Table 1 lists the model variables and parameters.

The LN-CTM model can be simplified for simulation of freeway traffic networks. While the general algorithm implements seperate Link and Node updates at each simulation step [5], the algorithm can be simplified to a four mode switching model for each link. For a section i (Figure 1), the density update equations belong to the following four modes - FF, CF, CC, and FC, for each link. Link i is updated using the CF mode equations, if Link i is in congestion and Link $i + 1$ is in freeflow. Other modes can be interpreted similarly. Here, Link i is classified to be in congestion if

Symbol	Name	Unit
	section length	miles
	period	hours
F_i	capacity	veh/period
v_i	free flow speed	section/period
w_i	congestion wave speed	section/period
n_i^J	critical density	veh/section
β_i	split ratio	dimensionless
k	period number	dimensionless
$f_i^{in}(k)$	flow into section i period k	veh/period
$f_i^{out}(k)$	flow out of section i period k	veh/period
$s_i(k), r_i(k)$	off-ramp, on-ramp flow in node i in period k	veh/period
$d_i(k)$	on-ramp demand for Link $i+1$ in period k	veh/period
$c_i(k)$	Total demand for Link $i+1$ in period k	veh/period

TABLE. 1 —— Model variables and parameters.

the flow into it is determined (limited) by its available capacity rather than total input demand, i.e. $c_{i-1} > \bar{w}_i(k)(n_i^J - n_i(k))$, where $c_{i-1}(k) = n_{i-1}(k)v_{i-1}(k)(1 - \beta_{i-1}(k)) + d_{i-1}(k)$ is the input demand into Link i and $\bar{w}_i(k)(n_i^J - n_i(k))$ represents the available output capacity. Here, $\bar{w}_i(k)$ and $\bar{v}_i(k)$ are defined as

$$\bar{w}_i(k) = min(w_i, \frac{F_i}{(n_i^J - n_i(k))})$$

$$\bar{v}_i(k) = min(v_i, \frac{F_i}{n_i(k)}) \tag{1}$$

The denisity update equations for the Links of the freeway can be summarised as,

(a) FF Mode

$$n_i(k+1) = n_i(k) + c_{i-1}(k) - n_i(k)\bar{v}_i(k) \tag{2}$$

(b) FC Mode

$$n_i(k+1) = n_i(k) + c_{i-1}(k) - \frac{\bar{w}_{i+1}(n_{i+1}^J - n_{i+1}(k))}{c_i(k)}n_i(k)\bar{v}_i(k) \tag{3}$$

(c) CC Mode

$$n_i(k+1) = n_i(k) + \bar{w}_i(n_i^J - n_i(k)) - \frac{\bar{w}_{i+1}(n_{i+1}^J - n_{i+1}(k))}{c_i(k)}n_i(k)\bar{v}_i(k) \tag{4}$$

(d) CF Mode

$$n_i(k+1) = n_i(k) + \bar{w}_i(n_i^J - n_i(k)) - n_i(k)\bar{v}_i(k) \tag{5}$$

The mainline flows can be determined by

$$f_i^{in}(k) = min(c_{i-1}(k), \bar{w}_i(k)(n_i^J - n_i(k)))$$

$$f_i^{out}(k) = \frac{min(c_i(k), \bar{w}_{i+1}(k)(n_{i+1}^J - n_{i+1}(k)))}{c_i(k)} n_i(k)\bar{v}_i(k) \tag{6}$$

while the offramp flows are deterimined by

$$s_i(k) = \beta_i(k)f_i^{out}(k) \tag{7}$$

The on-ramp flows and demands are given by

$$r_i(k) = \frac{min(c_i(k), \bar{w}_{i+1}(k)(n_{i+1}^J - n_{i+1}(k)))}{c_i(k)} d_i(k)$$

$$d_i(k+1) = d_i(k) + fl_i^{in}(k+1) - r_i(k) \tag{8}$$

where fl_i^{in} is the input flow for the onramp i.

3   IMPUTATION OF RAMP FLOWS

The LN-CTM model is utilized to impute the missing onramp input flows as well as the off-ramp split ratios for one day (24-hour) traffic flow simulation on a large freeway (eg. 40 miles) segment. The imputation procedure involves two stages - First, the total demands $c_i(k)$ are determined and then the demands and split-ratios are extracted from the total demand.

The imputation procedure employs an adaptive iterative learning procedure described in [7, 8]. It is assumed that the density and ramp flow profile is 24 hour periodic (i.e. the initial and final densities are assumed to be equal). This is not an restrictive assumption, since the freeway is found to be in free-flow (with low densities) around midnight. The LN-CTM algorithm is run multiple times, and at each run, the algorithm adapts the unknown demand estimates to minimize the error between the density generated by the model at each link and the data from the corresponding PeMS measurement. The procedure is repeated until the density error reduces to a sufficiently small value or stops decreasing.

As detailed in [7, 8],because of the 24 hour periodicity, the demand vector can be represented as a convolution of a kernel on a constant influence vector

$$c_i(k) = K(k)^T C_i \tag{9}$$

where $K(k)$ represent a 24 hour periodic time dependent kernel vector, and C_i is the influence vector. Some typical kernel functions ($K(k)$) include a unit-impulse or a gaussian window centered at time k.

The imputation procedure assumes initial estimates for the influence vectors $\hat{C}_i$. These estimates are then dynamically adapted at each time step, so that the model calculated densities, for all whole

freeway, match with the density profiles obtained from PeMS. At each time step, the mode for each cell is determined, and the corresponding learning update equations are used to adapt the influence vectors. The equations are given by

(a) FF Mode

$$\tilde{n}_i^o(k+1) = \hat{n}_i(k+1) - (\hat{n}_i(k) + K^T(k)\hat{C}_{i-1}(k) - \hat{n}_i(k)\hat{v}_i(k) - a(n_i(k) - \hat{n}_i(k)))$$

$$\tilde{n}_i(k+1) = \frac{\tilde{n}_i^o(k+1)}{1 + GK^T(k)K(k)}$$

$$\hat{C}_{i-1}(k+1) = \hat{C}_{i-1}(k) + GK(k)\tilde{n}_i(k+1)$$

$$\hat{n}_i(k+1) = \hat{n}_i(k) + K^T(k)\hat{C}_{i-1}(k+1) - \hat{n}_i(k)\hat{v}_i(k) - a(n_i(k) - \hat{n}_i(k)) \tag{10}$$

(b) FC Mode

$$\tilde{n}_i^o(k+1) = \hat{n}_i(k+1) - \left(\hat{n}_i(k) + K(k)^T\hat{C}_{i-1}(k) - \frac{\hat{w}_{i+1}(n_{i+1}^J - \hat{n}_{i+1}(k))}{K^T(k)\hat{C}_i(k)}\hat{n}_i(k)\hat{v}_i(k) - a(n_i(k) - \hat{n}_i(k)) \right)$$

$$\tilde{n}_i(k+1) = \frac{\tilde{n}_i^o(k+1)}{(1 + G'K^T(k)K(k) + GK^T(k)K(k))}$$

$$\hat{C}_{i-1}(k+1) = \hat{C}_{i-1}(k) + GK(k)\tilde{n}_i(k+1)$$

$$\hat{C}_i(k+1) = \hat{C}_i(k) - \frac{K(k)}{K^T(k)K(k)}\left(K^T(k)\hat{C}_i(k) - \frac{1}{1/K^T(k)\hat{C}_i(k) - G'K(k)\tilde{n}_i(k+1)} \right)$$

$$\hat{n}_i(k+1) = \hat{n}_i(k) + K(k)^T\hat{C}_{i-1}(k+1) - \frac{\hat{w}_{i+1}(n_{i+1}^J - \hat{n}_{i+1}(k))}{K^T(k)\hat{C}_i(k+1)}\hat{n}_i(k)\hat{v}_i(k) - a(n_i(k) - \hat{n}_i(k)) \tag{11}$$

(c) CC Mode

$$\tilde{n}_i^o(k+1) = \hat{n}_i(k+1) - \left(\hat{n}_i(k) + \hat{w}_i(n_i^J - \hat{n}_i(k)) - \frac{\hat{w}_{i+1}(n_{i+1}^J - \hat{n}_{i+1}(k))}{K^T(k)\hat{C}_i(k)}\hat{n}_i(k)\hat{v}_i(k) - a(n_i(k) - \hat{n}_i(k)) \right)$$

$$\tilde{n}_i(k+1) = \frac{\tilde{n}_i^o(k+1)}{(1 + G'K^T(k)K(k)}$$

$$\hat{C}_i(k+1) = \hat{C}_i(k) - \frac{K(k)}{K^T(k)K(k)}\left(K^T(k)\hat{C}_i(k) - \frac{1}{1/K^T(k)\hat{C}_i(k) - G'K(k)\tilde{n}_i(k+1)} \right)$$

$$\hat{n}_i(k+1) = \hat{n}_i(k) + \hat{w}_i(n_i^J - \hat{n}_i(k)) - \frac{\hat{w}_{i+1}(n_{i+1}^J - \hat{n}_{i+1}(k))}{K^T(k)\hat{C}_i(k+1)}\hat{n}_i(k)\hat{v}_i(k) - a(n_i(k) - \hat{n}_i(k)) \tag{12}$$

(d) CF Mode

$$\hat{n}_i(k+1) = \hat{n}_i(k) + \bar{w}_i(n_i^J - \hat{n}_i(k)) - \hat{n}_i(k)\hat{v}_i(k) \tag{13}$$

G and G' are positive gains. The parameter a is chosen so that the error equation is stable. The adaptation procedure is carried out through the entire density profile multiple times, so as to

reduce the 'error' $\sum |n_i(k) - \hat{n}_i(k)|$. Since the CF mode does not involve adaptation equations, the error may converge to a non-zero value for when this mode is in effect, while other modes shows negligible error. This occurs due to incorrect mode identification at that time instant. In this case, the corresponding estimates are "triggered" automatically so that the correct modes are identified. After the trigger, the adaptation procedure is continued, till the error becomes negligible or stops decreasing.

After determining the Total demand vector, the on-ramp demand and off-ramp split ratios are decoupled using a linear program. Figure 2 illustrates the position of the mainline detector, from which flow data is available. The linear program minimizes the objective $|(f_{i+1}^{in}(k) - f_{i+1}^{meas}(k)) - r_{i+1}(k)| + |(f_i^{out}(k) - f_{i+1}^{meas}(k)) - s_{i+1}(k)|$ to determine the missing on-ramp and off-ramp flows.
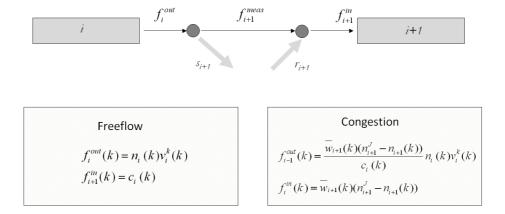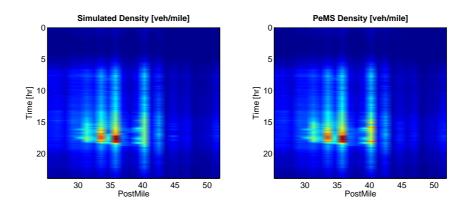

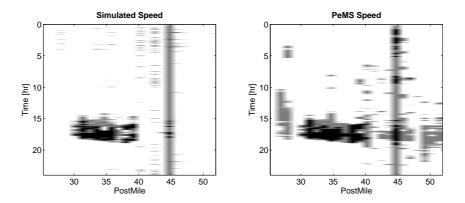
FIGURE. 2 — Decouple on-ramp and off-ramp flows.

4 APPLICATION

This section illustrates the application of the imputation algorithm to determine the on-ramp and off-ramp flow measurements in a 26-mile long section of I-210 E freeway in Pasadena. In this case, the freeway was divided into 26 links, and a total of 9 onramps and 5 offramps had either missing / incorrect data. The imputation procedure was carried out for these ramps. The fundamental diagram parameters for the links were obtained from an automated calibration procedure described in [9]. The final density error in the imputation was reduced to 2.63 %. Figure 3 shows that the density estimates have converged to their true values without appreciable error.
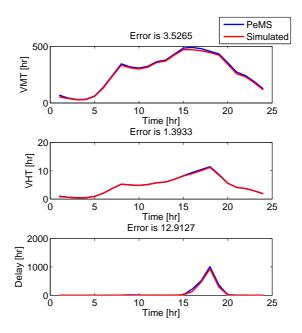
A simulation was performed with the imputed data. Figure 4 shows the simulated and the measured velocity contours, which show good agreement. The simulated contour plots clearly reproduce the locations of the major bottlenecks. The simulated and measured performance measures are compared in Figure 5, which also show good agreement. The simulated data had 2.63 % and 3.58 % density and flow errors respectively. Finally, Figure 6 lists the performance of the simulation as

FIGURE. 3 — Final density contours obtained after imputation.

FIGURE. 4 — Velocity Contours obtained from the I-210E simulation using imputed parameters.

FIGURE. 5 —— Performance measures - Vehicle Hours Travelled (VHT), Vehicle Miles Travelled (VMT) and Delay.

Test : for individual flows < 700	
total number of data points	80
number within 100 vph of data	70
85% of total number of data points	68
Test : for individual flows > 700 , < 2700	
total number of data points	141
number within 15% of data	116
85% of total number of data points	119
Test : for individual flows > 2700	
total number of data points	355
number within 400 vph of data	305
85% of total number of data points	301
Test : for sum of all link flows	
measured	2.13E+06
simulated	2.06E+06
Error	3.07%
Test : GEH for individual link flows	
total number of data points	576
number < 5	473
85% of total number of data points	489

FIGURE. 6 —— Final Density Contours obtained after imputation.

compared to the specifications provided in [10]. The simulation satisfies most of the requirements, while narrowly missing some of the criteria.

5 CONCLUSION

This paper has elaborated a novel imputation technique to determine missing ramp flow data. The imputation procedure was succesfully employed to determine the missing/incorrect onramp and offramp flows in a 26 mile portion of I210E freeway. The simulations, using the imputed on-ramp flows, and off-ramp split ratios, agree closely with the measurements, as shown by the velocity coutour plots, and performance measures plots and other calibration specification comparisons.

6 ACKNOWLEDGMENT

References

[1] California Department of Transportation, "Strategic Growth Plan," 2006. www.dot.ca.gov/docs/strategicgrowth.pdf - 2006-10-10.

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

[3] California Center for Innovative Transportation, "Corridor management plan demonstration: Final report," tech. rep., University of California, Berkeley, 2006. http://www.calccit.org/resources/2007_PDF/ CCIT_TO3_FinalReport-Jan5-07.pdf.

[4] PeMS, "PeMS website," 2007. http://pems.eecs.berkeley.edu, accessed 8/28/2007.

[5] A. Kurzhanskiy, *Modeling and Software Tools for Freeway Operational Planning.* PhD thesis, University of California, Berkeley, 2007.

[6] G. Gomes and R. Horowitz, "Optimal freeway ramp metering using the asymmetric cell transmission model," *Transportation Research, Part C*, vol. 14, no. 4, pp. 244–262, 2006.

[7] W. Messner, R. Horowitz, W.-W. Kao, and M. Boals, "A new adaptive learning rule," *IEEE Transactions on Automatic Control*, vol. 36-2, pp. 188–197, 1991.

[8] R. Horowitz, W. Messner, and J. B. Moore, "Exponential convergence of a learning controller for robot manipulators," *IEEE Transactions on Automatic Control*, vol. 36-7, pp. 890–892, 1991.

[9] G. Dervisoglu, G. Gomes, J. Kwon, and P. Varaiya, "Automatic calibration of the fundamental diagram and empirical observations on capacity." Submitted to TRB 88th Annual Meeting, 2009.

[10] Wisconsin Department of Transportation, "Freeway system operational assessment: Paramics calibration and validation," tech. rep., District 2, Milwaukee, 2002.