行政院國家科學委員會專題研究計畫　成果報告

即時因應事故衝擊之匝道儀控方法論研究

計畫類別：個別型計畫
計畫編號：NSC91-2211-E-009-046-
執行期間：91 年 08 月 01 日至 92 年 07 月 31 日
執行單位：國立交通大學交通運輸研究所

計畫主持人：許鉅秉

計畫參與人員：胡峻嘉, 陳宗鴻

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 92 年 9 月 8 日

A Stochastic Optimal Control based Approach to Real-Time Incident-Responsive Local Ramp Control

Jiuh-Biing Sheu[1], Yi-Hwa Chou[2], and Chi-Jen Yang[3]

***Abstract*—This paper presents a stochastic optimal control based methodology for real-time local ramp control in response to lane-blocking incidents on freeways. A discrete-time nonlinear stochastic system, which characterizes the time-varying relationships among lane traffic states, control variables, and point detector data, is proposed. In addition, a stochastic optimal control-based algorithm is developed to update the time-varying ramp metering control variables and incident-induced lane traffic states in real time. Preliminary test results indicate that the proposed ramp control method permits alleviating incident impacts efficiently particularly under low-volume and medium-volume incident conditions. Moreover, utilizing the proposed method, time-varying lane traffic states together with incident impacts on queue lengths and capacities can be estimated in parallel with incident-responsive ramp control to monitor the status of incident impact and provide incident-related traffic information for further applications, e.g., incident management.**

Keywords: Incident-responsive; Stochastic optimal control; Incident management

I. Introduction

Incident-induced traffic congestion remains to be a critical issue in the development of advanced freeway traffic management systems. Early literature [1, 2] has also pointed out that the growing

---

This work was supported by the National Science Council of Taiwan under Grant NSC 91-2211-E-009-046

[1] Associate Professor, Institute of Traffic and Transportation, National Chiao Tung University 4F, 114 Chung Hsiao W. Rd., Sec. 1 Taipei, Taiwan 10012 R.O.C. TEL: 886-2-2349-4963 FAX: 886-2-2349-4953 jbsheu@mail.nctu.edu.tw

[2] Professor, Department of Civil Engineering, National Taiwan University Taipei, Taiwan

[3] Research Assistant, Department of Civil Engineering National Taiwan University Taipei, Taiwan

incident-induced lane changes and queue lengths upstream to the incident site may significantly interrupt the traffic flows among adjacent lanes, leading to serious impacts on performance of the existing freeway traffic control and management systems.

Despite a variety of ramp control strategies that have been proposed [3, 4, 5, 6, 7, 8, 9, 10, 11, 12], there is still a lack of research on exploring real-time technologies to address the issues of time-varying incident impacts on ramp metering control. The investigations by Shaw [4] can be regarded as a pioneering study in incident-responsive ramp control, where they formulated the problem of traffic jams at the incident site with a deterministic queuing theory based model. Given the incident duration and some other ideal assumptions, Shaw et al. proposed to compare the predicted waiting time of any given on-ramp traffic arrival with a predetermined threshold to determine whether or not the new arrival is allowed to enter the freeway. The demand-capacity and percent-occupancy strategies [5, 6] which are the two typical ramp control modes used in the USA, and claimed to be incident-responsive, are based essentially on the same fundamental that the measurements such as volumes and occupancies upstream to the ramp under control are compared to preset thresholds to determine the ramp-metering rate. While under conditions of lane-blocking incidents, the determination in terms of these fixed thresholds may turn out to be a critical issue remaining in the aforementioned two traffic-responsive control algorithms because of the variety of incident-induced traffic flow patterns. To deal efficiently with the variety of freeway traffic congestion conditions including non-recurrent congestion cases Chen et al., [10] proposed an ingenious ramp control strategy employing fuzzy control theories. In comparison with the existing controllers at the study site under limited six incident scenarios, their test results implied that it is promising to gain higher ramp control efficiency with quick response to various incident cases via refined strategies. Wang [13] proposed to use a linear programming method together with a moving-average technique to formulate the problem of ramp metering control in response to nonrecurring congestion situations, where the ramp control rate was updated per five minutes in his approach to reduce incident impacts on mainline capacities of freeways.

Apparently, real-time incident-responsive ramp control warrants further research. Our early research [1] has pointed out that three levels of real-time functionality including (1) incident detection, (2) the prediction of incident congestion and (3) incident-responsive traffic management and control should be integrated to form a comprehensive incident management system, and herein, controlling ramp metering in real time to respond to the variety of incident-induced traffic congestion is a critical stage in the system development. Moreover, despite the significance of incident-responsive ramp control that has been recognized in some early literature noted above, limitations system competence in terms of real-time applications can be easily found in the published strategies. Most related issues stem from the lack of dedicated system functionality to characterize in real time incident impacts on traffic flows. For instance, the efficiency of the method by Wang [13] stated previously relies to a great extent on the accuracy in the estimation of aggregated vehicular travel time spent to go through the mainline detection zone while the effects of incident-induced lane traffic maneuvers such as lane changing and queuing may be ignored in this approach. Similar problems may also remain in the other ramp control strategies noted above.

This paper describes a new approach to real-time incident-responsive local ramp control. Under conditions of lane-blocking incidents, local ramp metering is formulated as a stochastic optimal control problem with the goal of minimizing the deviation between the estimated lane traffic states and the corresponding ideal values under incident conditions. The most distinctive feature of the proposed control method is that incident-induced inter-lane and intra-lane traffic states as well as incident impacts on traffic congestion can be estimated in real time, and then used as the parameters in the time-varying objective function to serve specific control purposes during lane-blocking incidents.

II. MODEL DESCRIPTION

The time-varying system states investigated herein consist of three groups, including (1) basic lane traffic states, (2) space-based incident impacts on traffic congestion, and (3) control variables. Specific detector configurations, as shown in Fig. 1, are proposed to collect raw traffic data that are used to update in real time the system states. Herein, the mainline segment bounded by the pair of detector stations refers to a control zone.

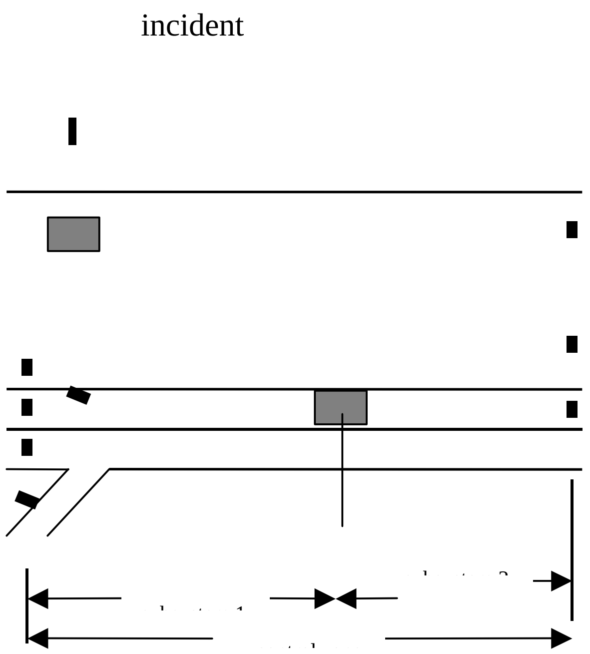

Fig. 1. Specification of detector configurations

In the following development, we consider a discrete-time, nonlinear stochastic model to formularize the relationships between the system states specified above under conditions of real-time ramp control in response to lane-blocking incidents on freeways. The proposed stochastic model is composed of four groups of dynamic equations, including (1) the objective function, (2) state equations, (3) measurement equations, and (4) state boundaries.

2.1 The Objective Function

Although there are diverse ramp control strategies proposed in early literature to optimize specific performance measures, e.g., delay, throughput, travel time, and speed, in this study we attempt to minimize the differences between the ideal and the estimated values of the basic lane traffic states. The ideal basic lane traffic states refer to the desired values of the basic lane traffic states that facilitate vehicular movement to the greatest extent under conditions of incident-induced traffic congestion, and herein, the upper bounds of the basic lane traffic states are considered to use. Therefore, we have the objective function ($\P$)

$$\P = \min E\left\{\sum_{k=0}^{N}[X(k)-\mathrm{I}]^T Q(k)[X(k)-\mathrm{I}]\right\} \quad (1)$$

where $X(k)$ is a $[6n_j + 2n_l]\times 1$ time-varying state vector containing the estimates of the time-varying basic lane traffic states; n_j and n_l represent the numbers of adjacent lanes and independent lanes, respectively; I is a $[6n_j + 2n_l]\times 1$ unit vector; $Q(k)$ represents a $[6n_j + 2n_l]\times[6n_j + 2n_l]$ time-varying diagonal, positive-definite weighting matrix; N corresponds to the total number of time steps in terms of incident duration. Coincidentally, the aforementioned ramp control fundamental tends to generate the results similar to that under the ramp control strategies serving the purposes of maximizing the mainline throughput in case of low-volume incidents and minimizing the total delay in case of high-volume incidents.

2.2 State Equations

The state equations denote the relationships of the basic lane traffic states in the temporal domain. To facilitate formulating the state equations of the discrete-time stochastic model, the assumption that the time-varying basic lane traffic states follow Gaussian-Markov processes is postulated. The rationalization of the assumption relies, to a certain extent, on our experiences in utilizing the concepts of random walk models to deal with the complexity of the patterns of system states changing in the presence of an incident [1, 2, 14]. That is, if there is no disturbance, the changing patterns of system states tend to be identical, and thus, easily to form a deterministic system; otherwise, the next-time-step system states may oscillate around the current-time-step system states by following Gaussian Processes. Note that the magnitude of the oscillation depends primarily on the pattern of the disturbance.

Based on the aforementioned postulation, we have the generalized form of the state equations given by:

$$X(k+1) = f[x(k)\,,\,k] + L[x(k)\,,\,\Omega(k)\,,\,k]w(k) \quad (2)$$

where $X(k+1)$ is a $[6n_j + 2n_l]\times 1$ time-varying vector of basic lane traffic states at time step $k+1$; $f[x(k), k]$ represents a deterministic term of the state equations with a $[6n_j + 2n_l]\times 1$ dimension of time-varying states estimated at time step k ($x(k)$); $L[x(k),\Omega(k),k]$ is a $[6n_j+2n_l]\times[6n_j+2n_l]$ state-dependent noise matrix, and in contrast, $w(k)$ corresponds to a $[6n_j+2n_l]\times 1$ state-independent Gaussian noise vector.

2.3 Measurement Equations

The measurement equations characterize the time-varying relationships between the measured lane traffic counts and the basic lane traffic states. Utilizing the specified relationships, the prior predictions of system states are updated in real-time in response to time-varying incident impacts under the proposed ramp control. The generalized form of the measurement equations can be expressed as:

$$Z(k) = h[x(k)\,,\,k] + v(k) \tag{3}$$

where $Z(k)$ is a $(n_i+n_j+n_l)\times 1$ time-varying measurement vector which is composed of the elements in terms of the lane traffic counts collected from the downstream detector station at time step k; similar to n_j and n_l defined previously, n_i corresponds to the number of blocked lanes; $h[x(k),k]$ is a $(n_i+n_j+n_l)\times 1$ time-varying vector in which each element associates a specific combination of the basic lane traffic states and measured lane traffic arrivals with a given element shown in $Z(k)$ to indicate the components of the measured downstream lane traffic counts. $v(k)$ is a $(n_i+n_j+n_l)\times 1$ Gaussian vector which represents the error terms of the collected traffic counts at time step k.

2.4 State Boundaries

The conditions of state boundaries are needed in the stochastic model to obtain feasible solutions of system state estimation. In the proposed model, three boundary conditions associated with the estimates of the basic lane traffic states (X(k)), the predictions of ramp control variables ($\Omega(k)$), and the consideration of the minimum ramp metering rate are formulated. Their generalized forms are depicted as follows:

$$\mathbf{0} \le X(k) \le \mathbf{1} \quad \text{for all } k \tag{4}$$

$$\mathbf{0} \le \Omega(k) \le 1 \quad \text{for all } k \tag{5}$$

$$\sum_{\pounds=0}^{n_-} [\Omega(k+\pounds)\times t] \ge T_{g,\min} \quad \text{for all } k \text{ and } - \tag{6}$$

where n_- is the maximum number of the sequential time steps which belong to a given cycle $-$; $T_{g,\min}$ represents the minimum on-ramp green time in a given cycle $-$.

III. CONTROL ALGORITHM

To perform the functionality of real-time incident-responsive traffic control utilizing the proposed stochastic model, a stochastic optimal control based algorithm is developed. The primary computational scenarios involved in the proposed algorithm include (1) system initialization, (2) prior prediction of system states, (3) stochastic optimal estimation of traffic states, and (4) determination of the time-varying ramp control variable. In order to obtain the minimum mean square estimates of the basic lane traffic states through the aforementioned scenarios (2) and (3), the fundamentals of an extended Kalman filter are applied. The concepts together with related application of the Kalman filtering techniques are stated elsewhere [14]. It is also worth mentioning that except the scenario of system initialization which is conducted only when the algorithm is triggered at the beginning of the ramp control period, the other three computational scenarios are executed in sequence at each time step until the beginning of the period in which incident impacts no longer exist. The control logic is summarized below.

Step 0. Initialize system states and the input raw traffic data. Given $k=0$, system states including (1) the basic lane traffic states $X(0|0)$, (2) the covariance matrix of the state estimation error $\Phi(0|0)$, and (3) the weighting matrix $Q(0)$ are initialized. In addition, let the time-varying ramp control variable $\Omega(0)=0$ to ensure that the on-ramp vehicles do not contribute durably to deterioration on the incident impacts at the onset of the incident.

Step 1. Compute prior predictions of lane traffic state variables ($X(k+1|k)$) and the covariance matrix of the state estimation error ($\Phi(k+1|k)$), respectively, by:

$$X(k+1|k)=f[x(k),\Omega(k),k] \tag{7}$$

$$\Phi(k+1|k)=F(k)\Phi(k|k)F^T(k)+L[x(k),\Omega(k),k]Q(k)L[x(k),\Omega(k),k]^T \tag{8}$$

where $F^T(k)$ is the transpose matrix of $F(k)$.

Step 2. Calculate the Kalman gain by:

$$K(k+1)=\Phi(k+1|k)H^T(k+1)\left[H(k+1)\Phi(k+1|k)H^T(k+1)+R(k+1)\right]^{-1} \tag{9}$$

where $R(k+1)$ is the covariance matrix of $v(k+1)$; and $H(k+1)$ is denoted by:

$$H(k+1)=\frac{\partial h[x(k+1),k+1]}{\partial X(k+1)}\Bigg|_{x(k+1)=x(k+1|k)} \tag{10}$$

Step 3. Update the prior estimates of the basic lane traffic states ($X(k+1|k+1)$) by:

$$X(k+1|k+1)=X(k+1|k)+K(k+1)\Delta Z(k+1|k) \tag{11}$$

where $\Delta Z(k+1|k)$ is given by:

$$\Delta Z(k+1|k)=Z(k+1)-h[x(k+1|k),k+1] \tag{12}$$

Step 4. Truncate the estimates of the basic lane traffic states variables ($X(k+1|k+1)$) with the conditions of state boundaries, and normalize incident-induced mandatory lane-changing fractions in a given blocked lane such that:

$$\sum_{\forall j \in J} p_{ij}^{1}(k+1) \quad \leq \quad 1$$

(13)

Step 5. Update the covariance matrix of the state estimation error ($\Phi(k+1\,|\,k+1)$) as:

$$\Phi(k+1\,|\,k+1) = [I - K(k+1)H(k+1)]\Phi(k+1\,|\,k) \qquad (14)$$

Step 6. Update the states of the space-based incident impacts at the end of time step k+1. In this step, the formulae of the space-based incident impacts specified previously are employed.

Step 7. Calculate the decision-variable vector $\Omega(k+1)$. According to the fundamentals of stochastic optimal control theories, the estimates of the basic lane traffic states ($X(k+1\,|\,k+1)$) are fed back through the optimal gain matrix $E(k+1)$ to achieve the goal of the pre-specified objective function by:

$$\Omega(k+1) = -E(k+1)X(k+1\,|\,k+1) + \text{¦}(k+1) \qquad (15)$$

Herein, $E(k+1)$ and $\text{¦}(k+1)$ are denoted respectively by:

$$E(k+1) = \left[B^{T}(k+1)S(k+2)B(k+1) + R(k+1)\right]^{-1} B^{T}(k+1)S(k+2)F(k+1) \qquad (16)$$

$$\text{¦}(k+1) = \left[B^{T}(k+1)S(k+2)B(k+1) + R(k+1)\right]^{-1}\left[B(k+1)Q(k+1)X(k+1) + R(k+1)\Omega(k+1)\right] \qquad (17)$$

where the matrix $S(k+2)$ should satisfy the Riccati equation.

Step 8. Check the estimate of the time-varying ramp control variable to satisfy both the conditions of state boundaries and the minimum ramp metering rate shown in Eqs. (5) and (6), respectively.

Step 9. Check incident status by conducting the following rules:

If the incident is removed and the queues in blocked lanes no longer exist, then stop the ramp control algorithm. Otherwise, input the next-time-step raw traffic data; let the time step index k=k+1, and then go back to **Step 1** to continue the control algorithm.

IV. Preliminary Tests

Considering the difficulty in gathering enough real incident-related traffic data for diverse incident cases, simulation data generated from the Paramics microscopic traffic simulator was used in the test scenario. The Paramics simulator was calibrated prior to this study, and tasks related to evaluating, qualitatively and quantitatively, the Paramics simulator can also be found in our early related research [15]. To simulate diverse lane-blocking incidents in a given detection zone on a freeway, a simplified 3-lane freeway segment which is 3 km in length, and comprises one signalized on-ramp was built via Paramics. Lane-blocking incidents were mainly generated on the mainline segment of the study site, and then the output data simulated from Paramics were collected at each 10-sec. time step. In this test scenario, twenty-seven incident cases associated with diverse incident attributes including incident duration, incident location in a given lane, the lane blocked, and traffic flow condition were simulated.

Given incident attributes diverse lane-blocking incidents under control of two specific ramp control strategies including the proposed control method and pretimed control mode were simulated through Paramics, respectively. Measures including lane traffic loads and queue length were gathered for the use in comparison of the system performance. Details about input data generation and parameter calibration are described elsewhere [16]. Table 1 lists the numerical comparison results.

Table 1. Comparison of ramp control performance

Overall, the results summarized in Table 1 reveal a certain improvement in ramp control performance made by the proposed control method in contrast with the pretimed control strategy. By comparing the results of the high-volume case, the system performance in terms of reductions in space-based incident impacts is improved by utilizing the proposed control method. Such a consequence is not surprising because as explicated in formulating the objective function of the proposed stochastic model, the proposed ramp control strategy aims at minimizing the incident impacts. The aforementioned argument is evidenced particularly in both medium-volume and low-volume incident cases. Overall, the improvement in control performance by the proposed method turns out to be relatively more significant in these cases, compared to the high-volume case.

V. Concluding Remarks

This paper has presented a stochastic optimal control based method for real-time local ramp control under conditions of lane-blocking incidents on mainline segments of freeways. To achieve the greatest reduction of incident impacts on traffic congestion in real-time via stochastic optimal control-based technologies, we specified three groups of time-varying lane traffic variables, and then proposed a discrete-time nonlinear stochastic model as well as a real-time ramp control algorithm.

Our numerical results revealed the applicability of the proposed local ramp control method in terms of responding to incident-induced traffic congestion as well as estimating system states in real time under conditions of various incident cases on freeways. Results presented here also suggested the relative advantages of the proposed control method compared with two other specified ramp control strategies. More importantly, the proposed approach may indicate its potential in terms of characterizing incident-induced lane traffic states together with incident impacts in real time in the procedure of real-time incident-responsive control

Nevertheless, more tests as well as comparisons with other advanced ramp control strategies warrant further research to verify the robustness of the proposed incident-responsive ramp control method. The extension of the proposed approach as well as system scope for the scenarios of corridor control will be undertaken in our further research. Moreover, efforts on either integrating the proposed ramp control method with other advanced traffic control and management technologies including variable message signs (VMS) seem to be needed for high-volume incident cases.

REFERENCES

[1] J.-B. Sheu, Y.-H. Chou, and L.-J Shen, "A stochastic estimation approach to real-time prediction of incident effects on freeway traffic congestion," *Transportation Research-Part B*, vol. 35B, no. 6, pp. 575-592, 2001.

[2] J.-B. Sheu and S. G. Ritchie, "Stochastic modeling and real-time prediction of vehicular lane-changing behavior," *Transportation Research-Part B*, vol. 35B, no. 7, pp. 695-716, 2001.

[3] J. A. Wattleworth and D. S. Berry, "Peak-period control of a freeway system-some theoretical investigations," *Highway Research Record* vol. 89, pp.1-25, 1965.

[4] L. Shaw and W. R. McShane, "Optimal ramp control for incident response," *Transportation Research* vol. 7, pp. 393-411, 1973.

[5] D. P. Masher, D. W. Ross, P. J. Wong, P. L. Tuan, Zeidler, and S. Peracek, "Guidelines for design and operating of ramp control systems," Standford research Institute Report NCHRP 3-22, SRI Project 3340. SRI, Menid Park, CA, 1975.

[6] H. M. Koble, T. A. Adams, and V. S. Samant, "Control strategies in response to freeway incidents," Federal Highway Administration, Washington D. C, Report No. FHWA/RD-80/005, 1980.

[7] M. Papageorgiou, H. S. Habib, and J. M. Blosseville, "ALINEA: a local feedback control law for on-ramp metering," *Transportation Research Record* vol. 1320, 58-64, 1991.

[8] H. M. Zhang and S. G. Ritchie, "Freeway ramp metering using arterial neural networks," Transportation Research Part C, vol. 5C, no. 5, pp. 273-286, 1997.

[9] G. L. Chang, P. K. Ho and C. H. Wei., "A dynamic system-optimum control model for commuting traffic corridors," *Transportation Research Part C*, vol. 1, no. 1, pp. 3-32, 1993.

[10] L. L. Chen, A. D. May, and D. M. Auslander, "Freeway ramp control using fuzzy set theory for inexact reasoning," *Transportation Research. Part A*, vol. 24, no. 1, pp. 15-25, 1990.

[11] M. Papageorgiou, "An integrated control approach for traffic corridors," *Transportation Research Part C*, vol. 3, no. 1, pp. 19-30, 1995.

[12] H. Zhang and W. W. Recker, "On optimal freeway ramp control policies for congested traffic corridors," Transportation Research Part B, vol. 33, pp. 417-436, 1999.

[13] M.-H. Wang, "Optimal ramp metering policies for nonrecurring congestion with uncertain incident duration," Ph.D. thesis, Purdue University, USA, 1994.

[14] J.-B. Sheu, "Stochastic estimation of lane-changing probabilities and its application to incident detection," Ph.D. Thesis, University of California, Irvine, USA, 1997.

[15] B. Abdulhai, J.-B. Sheu, and W. W. Recker, “Simulation of ITS on the Irvine FOT area using the 'Paramics 1.5' scalable microscopic traffic simulator: phase I : model calibration and validation,” California PATH research report UCB-ITS-PRR-99-12 1055-1425, 1999.

[16] C.-J. Yang, “Prototype of a ramp control strategy for alleviating freeway incident congestion,” Master Thesis, Dept. Civil Eng., National Taiwan Univ., Taipei, Taiwan, 2001.

Table 1. Comparison of ramp control performance

incident case	control strategy / evaluation measure	proposed method (mode-1)	pretimed control (mode-2)	relative improvement Mode-1 vs. mode-2
high-volume	lane traffic loads (veh/10-sec)	58	61	5.2%
	queue length (veh/10-sec)	86	91	5.9%
medium-volume	lane traffic loads (veh/10-sec)	30	38	26.7%
	queue length (veh/10-sec)	28	42	50.0%
low-volume	lane traffic loads (veh/10-sec)	4	5	25%
	queue length (veh/10-sec)	6	7	1.7%