A STUDY ON THE DYNAMIC RAMP CONTROL STRATEGY UNDER CONSIDERATION OF SURFACE STREETS

Kotaro Kumagai
Engineering Division, Kumagai Gumi Co., LTD.
2-1, Tsukudo-cho, Shinjuku-ku, Tokyo, 162-5526, Japan
Tel: (81) 3 5261 5526 - Fax (81) 3 5261 9350 - e-mail: kkumagai@ku.kumagaigumi.co.jp

Masao Kuwahara
Institute of Industrial Science, University of Tokyo
4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, Japan
Tel: (81) 3 5452 6231 - Fax (81) 3 5452 6418 - e-mail: kuwahara@nishi.iis.u-tokyo.ac.jp

Toshio Yoshii
Department of Infrastructure Systems Engineering, Kochi University of Technology
185, Miyanokuchi, Tosayamada, Kochi, 782-8502, Japan
Tel: (81) 3 887 57 2406 - Fax (81) 3 887 57 2420 - e-mail: yoshii@infra.kochi-tech.ac.jp

SUMMARY

This study analyses theoretical flow patterns of the Dynamic System Optimum (DSO) assignment and proposes the ramp control strategy toward DSO using a simple network of a parallel pair of an expressway and a surface street. The marginal cost for the dynamic traffic flow is first defined in contrast with one in static analysis. An analysis on DSO based upon the dynamic marginal cost suggests that the best strategy for DSO is to assign demand onto the faster route (the expressway) just up to its capacity. Then, assuming each traveler chooses a route so as to minimize his/her own travel time, we show that a ramp control such as metering can decrease the total travel time of not only the expressway but also the surface street. Finally, we apply the proposed ramp control strategy to the existing traffic conditions.

DSO ASSIGNMENT

Network and Demand

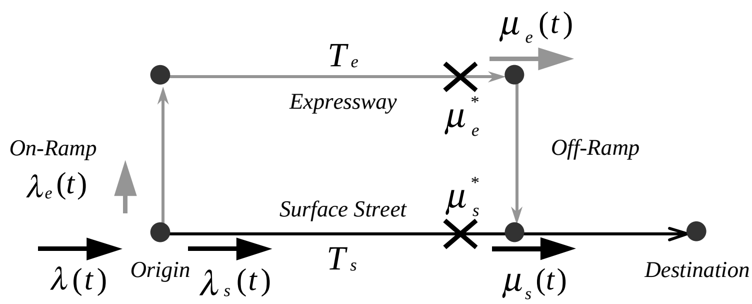

Figure 1 A Study Network with Single OD Pair

Figure1 shows a study network consisting of an expressway and a surface street. Such a simple network is used to focus on the key issue. A time-dependent traffic demand from single origin to single destination given. We suppose the following three assumptions.

1. Free flow travel time of the expressway T_e is less than one of the surface street T_s.
2. A queue has no physical length.
3. Each vehicle is served in the FIFO (First-In-First-Out) discipline at the both bottlenecks.

The OD demand rate departing from the origin is written as $\lambda(t)$, which is assumed to have a single peak and is split onto the expressway and the surface street. The corresponding arrival and departure rate are written as below.

$\lambda_e(t)$, $\lambda_s(t)$: Arrival rates entering the expressway and the surface street at time t.

$\mu_e(t)$, $\mu_s(t)$: Departure rates leaving the expressway and the surface street at time t.

And their cumulative functions are defined as below.

$A_e(t)$, $A_s(t)$: The cumulative number of vehicles entering the expressway and the surface street by time t, $A(t)=A_e(t)+A_s(t)$.

$D_e(t)$, $D_s(t)$: The cumulative number of vehicles leaving the expressway and the surface street by time t.

The μ^*_e and μ^*_s in Figure 1 mean capacities of the expressway and the surface street, respectively.

DYNAMIC MARGINAL COST

The total travel time TC during the study time period, $0 \le t \le \tau$, is written as below.

$$TC = TC_e + TC_s = \int_0^\tau \{T_e + w_e(t)\} \cdot \lambda_e(t)dt + \int_0^\tau \{T_s + w_s(t)\} \cdot \lambda_s(t)dt \qquad (1)$$

where,

TC_e : Total travel time of the expressway.

TC_s : Total travel time of the surface street.

$w_e(t)$: Waiting time of a vehicle entering the expressway at time t.

$w_s(t)$: Waiting time of a vehicle entering the surface street at time t.

Let us consider the marginal cost on the expressway and the surface street, $MC_e(t)$ and $MC_s(t)$, which describe how much travel time changes when unit arrival rate at time t shifts. For the expressway,

$$\begin{aligned}
MC_e(t) &= \frac{dTC_e}{d\lambda_e(t)dt} = \frac{d\int_0^\tau \{T_e + w_e(u)\} \cdot \lambda_e(u)du}{d\lambda_e(t)} \cdot \frac{1}{dt} \\
&= T_e + w_e(t) + \int_t^\tau \frac{dw_e(u)}{d\lambda_e(t)} \lambda_e(u)du \cdot \frac{1}{dt} \\
&= T_e + w_e(t) + \int_t^\tau \frac{dt}{\mu_e^*} \lambda_e(u)du \cdot \frac{1}{dt} \\
&= T_e + w_e(t) + \frac{1}{\mu_e^*}(A_e(t_1^e) - A_e(t)) \\
&= T_e + w_e(t) + \{(t_1^e + T_e) - (t + T_e + w_e(t))\} \\
&= T_e + t_1^e - t
\end{aligned} \qquad (2)$$

where t_1^e : queue vanishing time on the expressway (see Figure 2)

Equation (2) implies that, in contrast with the static marginal cost, shift of arrival rate at time t affects travel time of all vehicles arriving at the bottleneck from t until the queue vanishes. That is, since the cumulative number from time t increases by one unit, the marginal cost is interpreted as $MC_e(t)$ = travel time of the vehicle entering at time t ($T_e + w_e(t)$) + total waiting time change of vehicles entering after t ($(A_e(t_1^e) - A_e(t))/\mu_e^*$). Similarly, for the surface street, $MC_s(t) = T_s + t_1^s - t$. To establish DSO, we should clearly equilibrate these dynamic marginal costs on both routes.

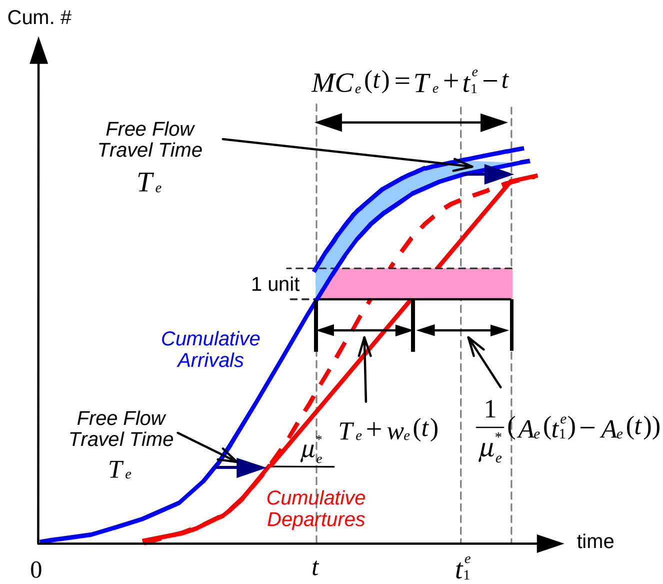

Figure 2 Dynamic Marginal Cost on the Expressway

Strategy for DSO

At first, we consider a case where no queue forms on the surface street because of the sufficient capacity (μ^*_e<Max $\lambda(t)$<$\mu^*_e+\mu^*_s$). When free flow travel time T_e is equal to T_s, the solution of DSO must be the same as one in DUE. Since users have no preference one route to another one in this case, the problem can be reduced to a single bottleneck with capacity of $\mu^*_e+\mu^*_s$.

When $\Delta T = T_s - T_e > 0$, the solution of DUE would be one shown in the left upper figure of Figure 3. For simplicity, we assume free flow travel time on the expressway T_e is equal to zero thereafter. The $A(t)$ shows the total cumulative demand, all of which uses the expressway at the beginning. At time t_0, a queue starts forming on the expressway and the travel time on the expressway becomes equal to T_s at time t_2. From t_2 to t_3, the demand is split so as to equalize travel times on both routes. However, at time t_3, the whole demand rate decreases lower than expressway capacity, and then the entire demand returns to the expressway. Thus, the slope of $A_e(t)$ for the expressway becomes equal to that of total demand $A(t)$ after time t_3. The left lower figure of Figure 3 shows the marginal costs of both routes. When a queue forms on the expressway at time t_0, $MC_e(t)$ suddenly jumps to $t_1 - t_0 + T_e$ and then decreases until time t_1 with slope of –1 according to equation (2). On the other hand, the marginal cost of the surface street $MC_s(t)$ stays constant value of T_s. From this figure, in DUE, the marginal costs are not in the equilibrium; that is, more demand tends to be assigned to the expressway especially from time t_0 where $MC_e(t)>MC_s(t)$.

The right upper figure of Figure 3 shows the queue evolution under DSO. Until time t_0, both marginal costs stay to their free flow travel times because of no queues, and the entire demand uses the expressway. At time t_0, the whole demand rate becomes equal to expressway capacity μ^*_e. Under this condition where $\lambda_e(t)=\mu^*_e$, if we add one unit of demand onto the expressway, a queue on the expressway would last until time t_1. Therefore, the marginal cost of one unit addition, $MC_e^+(t)$, is $t_1 - t_0+T_e$. On the other hand, if we subtract one unit of demand, demand rate on the expressway $\lambda_e(t)$ is less than its capacity μ^*_e and $MC_e^-(t)$ is T_e ($<T_s$) as shown in the broken line in the right lower Figure 3. As seen here, the marginal cost jumps between $MC_e^+(t)$ and $MC_e^-(t)$ when demand of just equal to its capacity is assigned. Hence, if we add one more unit on the expressway, the expressway marginal cost $MC_e^+(t)$ becomes larger than the surface street marginal cost $MC_s(t)$, while if one unit is subtracted, $MC_e^-(t)$ becomes smaller than $MC_s(t)$. This suggests, from t_0 to t_4, to assign the demand onto the expressway just up to its capacity to keep the bottleneck busy but not more than its capacity. From time t_0, $MC_e^+(t)$ keeps decreasing with slope of –1 until time t_4. After time t_4, since the expressway marginal cost $MC_e^+(t)$ becomes smaller than the surface street marginal cost $MC_s(t)=T_s$, the entire demand should be assigned to the expressway. Thus, if the total demand rate is larger than its capacity μ^*_e for t_4 to t_1, a queue forms on the expressway as shown in the right upper Figure 3. It is interesting that $t_1 - t_4$ must be equal to ΔT, since $MC_e^+(t)$ still decreases with slope of –1 from t_4 to t_1. To draw Figure 3, t_0 is first determined so that $A(t_0)=\mu^*_e$. Then, a straight line with slope μ^*_e is superimposed with the total demand curve $A(t)$ and find two intersection points t_4 and t_1 so that $t_1 - t_4 =\Delta T$. If ΔT gets longer, $t_1 - t_4$ is therefore increases. For sufficient large ΔT, t_4 becomes equal to t_0. Then, a time interval between t_0 and t_4 during which some demand uses the surface street is disappeared and hence the entire demand uses the expressway all the time. In other words, no one should take the surface street for DSO because of its quite long travel time T_s.

If Max $\lambda(t)>\mu^*_e+\mu^*_s$ then, a queue forms both on the expressway and on the surface street. For all cases, a basic strategy to establish DSO is to assign demand onto the faster route (the expressway) just up to its capacity as discussed above. This strategy would be valid not only a case with only two alternative routes as we discussed but also cases with more than two routes. That is, we should assign demand first onto the fastest route up to its capacity and then assign the remainder to the second best also up to the capacity and so on.

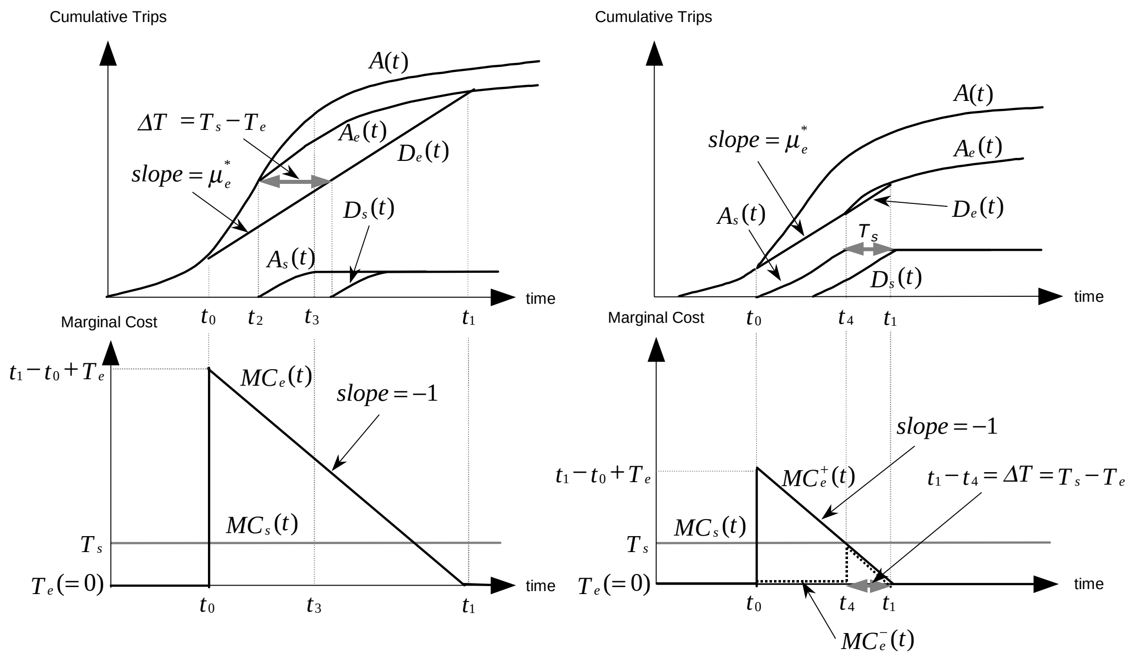

Figure 3 Queue Evolution in DUE(left) and DSO(right)
(No Queue on Surface Street)

RAMP CONTROL

We have discussed on DSO, in which we directly control demand, itself. However, in reality, we cannot order drivers which way to go. They can be controlled only through some physical and/or economical means such as ramp control, road pricing, and so on. In this section, let us consider the ramp control that restricts the flow rate entering an expressway. We also assume that traffic condition becomes DUE without ramp control. Thus, the point of the analysis is whether the ramp control reduces the total travel time compared with the DUE condition.

Preliminary Consideration

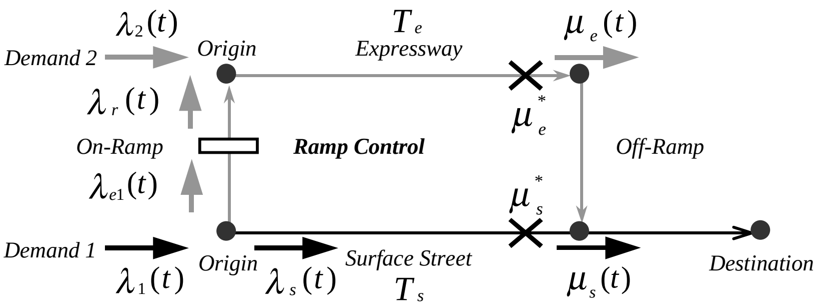

Figure 4 A Study Network with Two OD Pairs

This section summarizes conditions where the ramp control is or is not clearly effective to reduce the total travel time. First, let us consider a single OD situation where *Demand 2* in Figure 4 is equal to 0. For this single OD, travel time reduction is not possible because of the following reason. If you reduce the on-ramp inflow rate above the expressway bottleneck capacity, the true bottleneck on the expressway route is still at the expressway bottleneck. Thus the situation would be the same as one without ramp control. On the other hand, if you reduce the inflow rate below the expressway bottleneck capacity, you simply shift the bottleneck from the expressway to the on-ramp, and the total travel time would increase. Consequently, if we have only single OD, the ramp control is not effective.

Let us add one more OD (*Demand 2*), whose rate is $\lambda_2(t)$ to the same destination as shown in Figure 4. The *Demand 2* has only the expressway route while *Demand 1* has two alternative routes. With this additional OD *Demand 2*, the ramp control would be effective in some situations. When a queue forms on the expressway, the bottleneck capacity of μ^*_e is shared by *Demands 1* and *2*. Thus, if you control the inflow rate, you can control the share of μ^*_e between *Demands 1* and *2*. For instance, since the bottleneck service is FIFO, if you reduce the inflow rate of *Demand 1* to the expressway, the share of *Demand 2* in μ^*_e would increase, and vice versa. This is the role of the ramp control. When no queue forms on the surface street due to the sufficient capacity, travel time of *Demand 1* does not get larger than free flow travel time on the surface street, T_s, independent of the ramp control. This means that, for *Demand 1*, their travel time is the same as one in the DUE condition whatever ramp control is implemented. Therefore, apparently we had better restrict the inflow rate of *Demand 1* to the expressway so that *Demand 2* can pass the expressway bottleneck without any delay.

Another obvious case is that if the maximum flow rate of $\lambda(t)=\lambda_1(t)+\lambda_2(t)$ is smaller than or equal to expressway capacity μ^*_e, the whole demand can use the expressway without delay. This means the ramp control has no effect.

Effects of Ramp Control

This section quantitatively analyzes effects of the ramp control under $\Delta T = T_s - T_e > 0$, $\lambda_2(t) > 0$, and Max $\lambda(t) > \mu^*_e$. The $\lambda_1(t)$, is divided into $\lambda_{e1}(t)$ for the expressway and $\lambda_s(t)$ for the surface street as shown in Figure 4. The $\lambda_{e1}(t)$ is restricted at the ramp and only rate of $\lambda_r(t)$ can enter the expressway. Thus, if $\lambda_r(t) < \lambda_{e1}(t)$, a queue forms at the ramp. The $\lambda_r(t)$ is our control variable. Figure 5 shows cumulative curves under some ramp control $\lambda_r(t)$, in which we find three different time period.

In time period 1 from T_0 to T_1, all the demand uses the expressway because of no delay on the expressway at the beginning. But from time T_0, a queue start growing on the expressway and the travel time becomes equal to one on the surface street at time T_1. For $T_0 \leq t \leq T_1$, the departure rate from the bottleneck is therefore equal to expressway capacity μ^*_e. Arrival rates of *Demand 1* and *2* at the expressway bottleneck are $\lambda_r(t)$ and $\lambda_2(t)$ respectively during the period. Based on FIFO, the capacity μ^*_e is split into $\mu_{e1}(t)$ for *Demand 1* and $\mu_{e2}(t)$ for *Demand 2* as follows:

$$\mu_{e1}(t + w_e(t)) = \mu^*_e \frac{\lambda_r(t)}{\lambda_2(t) + \lambda_r(t)}, \quad \mu_{e2}(t + w_e(t)) = \mu^*_e \frac{\lambda_2(t)}{\lambda_2(t) + \lambda_r(t)} \tag{3}$$

Without ramp control, ratio $\mu_{e1}(t+w_e(t))/\mu_{e2}(t+w_e(t))$ is always equal to $\lambda_{e1}(t)/\lambda_2(t)$ but here $\mu_{e1}(t)$ and $\mu_{e2}(t)$ can be controlled by $\lambda_r(t)$ as in equations (3). On the other hand, to make the ramp control effective, the ramp itself should not be a bottleneck:

$$\lambda_r(t) \geq \mu_{e1}(t + w_e(t)) = \mu^*_e \frac{\lambda_r(t)}{\lambda_2(t) + \lambda_r(t)} \;\Leftrightarrow\; \lambda_r(t) \geq \mu^*_e - \lambda_2(t) \tag{4}$$

Also, arrival rate at the ramp $\lambda_{e1}(t)$ should be larger than the control rate $\lambda_r(t)$:

$$\lambda_r(t) \leq \lambda_{e1}(t) \tag{5}$$

Since everyone uses the expressway during this period ($\lambda_{e1}(t) = \lambda_1(t)$), the constraints (4) and (5) are summarized as

$$\mu^*_e - \lambda_2(t) \leq \lambda_r(t) \leq \lambda_1(t) \tag{6}$$

From this, clearly $\lambda_1(t)+\lambda_2(t)$ should be greater than μ^*_e, which suggests that the ramp control should obviously start after time when the total demand rate of $\lambda_1(t)+\lambda_2(t)$ exceeds μ^*_e. If constraint (4) is satisfied, the essential bottleneck for *Demand 1* is still at the expressway bottleneck (not the on-ramp). Thus, under constraint (4), we can evaluate whole delay of *Demand 1* at the expressway bottleneck from $\lambda_1(t)$ and $\mu_{e1}(t)$, which is controlled by $\lambda_r(t)$ as in equations (3).

In time period 2 from T_1 to T_2, a vehicle departing the origin at time $T_1 - \Delta T$ faces the same travel time on the two alternative routes. From this time, *Demand 1* also uses the surface street and is split into $\lambda_s(t)$ and $\lambda_{e1}(t)$ so as to equalize these travel times of *Demand 1*. Even in this period, $\mu_{e1}(t)$ and $\mu_{e2}(t)$ is obtained from equations (3), and constraints (4) and (5) must be satisfied.

In time period 3 from T_2 to T_3, travel time on the expressway becomes faster and *Demand 1* also uses the expressway. The entire departure rate from the bottleneck is thus μ^*_e.

To draw Figure 5, first, cumulative curves $A_1(t)$, $A_2(t)$ and $A_1(t)+A_2(t)$ are drawn with given slopes $\lambda_1(t)$, $\lambda_2(t)$ and $\lambda_1(t)+\lambda_2(t)$ respectively. Second, from time T_0 when $\lambda_1(t)+\lambda_2(t)$ becomes equal to expressway capacity μ^*_e, a straight line with slope of μ^*_e are drawn. Third, inflow rate from the ramp $\lambda_r(t)$ is also drawn based upon our control strategy. Fourth, cumulative curve with slope of $\lambda_r(t)+\lambda_2(t)$ is drawn, which is the flow rate arriving at the expressway bottleneck. Since the bottleneck service is FIFO, $\mu_{e1}(t)$ and $\mu_{e2}(t)$ can be determined from equations (3). Fifth, from time $T_1 - \Delta T$, since users start using also the surface street, the departure curve from the surface street bottleneck with slope of μ^*_s is drawn.

Sixth, arrival curves of *Demand 1* at the expressway and the surface street bottlenecks are determined so as to establish the equilibrium; that is,

$$\frac{\lambda_1(t)}{\mu_{e1}(t+w_e(t))+\mu_s^*}=\frac{\lambda_{1e}(t)}{\mu_{e1}(t+w_e(t))}=\frac{\lambda_s(t)}{\mu_s^*} \tag{7}$$

At time T_2 - ΔT, travel time on the expressway returns to ΔT, and no one would like to use the surface street thereafter. Therefore, arrival rate of *Demand 1* at the expressway $\lambda_{e1}(t)$ is equal to $\lambda_1(t)$ from time T_2 - ΔT.

As explained above, it is interesting that the total departure rate from the both bottlenecks $\mu(t)$ during each of the three time periods is written independent of control variable $\lambda_r(t)$:

$$\mu(t) = \begin{cases} \mu^*_e & \text{for} \quad T_0 \le t < T_1 \\ \mu^*_e + \mu^*_s & \text{for} \quad T_1 \le t < T_2 \\ \mu^*_e & \text{for} \quad T_2 \le t < T_3 \end{cases} \tag{8}$$

Hence, to minimize the total travel time, we should enlarge time period 2 in which the total departure rate is the largest. First, to let T_1 be earlier, $\mu_{e1}(t)$ should be smaller so that delay of *Demand 1* quickly becomes equal to ΔT. From equations (3), $\mu_{e1}(t)$ is written as a function of $\lambda_r(t)$ and $d\mu_{e1}(t+w_e(t))/d\lambda_r(t) \geq 0$. Thus, we should control $\lambda_r(t)$ as small as we can under constraints of (4) and (5). Clearly from (4), $\lambda_r(t)=\mu^*_e - \lambda_2(t)$. For the extreme case, if $\mu^*_e - \lambda_2(t)<0$, the ramp should be closed: $\lambda_r(t)=0$. When the ramp is closed, T_1 - T_0 is equal to ΔT and T_1 is the smallest. Second, to let T_2 be larger, we should again control $\mu_{e1}(t)$ as small as possible so that delay of *Demand 1* return to ΔT slowly. Similarly, the optimum control strategy is to adjust $\lambda_r(t)$ to be $\mu^*_e - \lambda_2(t)$. In conclusion, the optimum control strategy is to keep $\lambda_r(t)$ being $\mu^*_e - \lambda_2(t)$ so that the expressway capacity is assigned to *Demand 2* as possible and only remaining capacity is assigned for *Demand 1*. This strategy is consistent with our discussion on DSO in the previous sections.

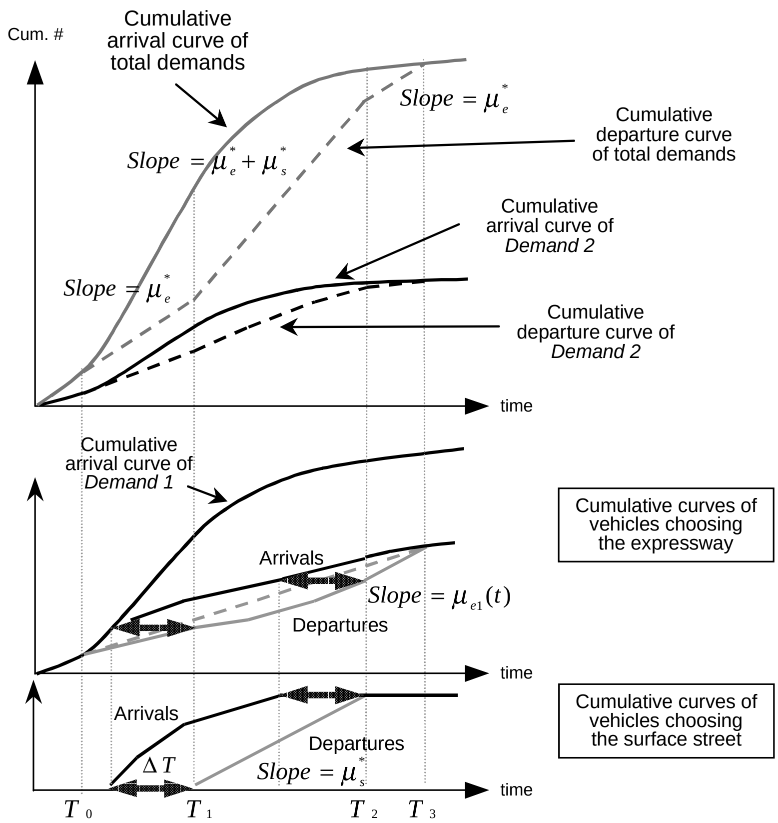

Figure 5 Queue Evolution with Ramp Control

Case Study

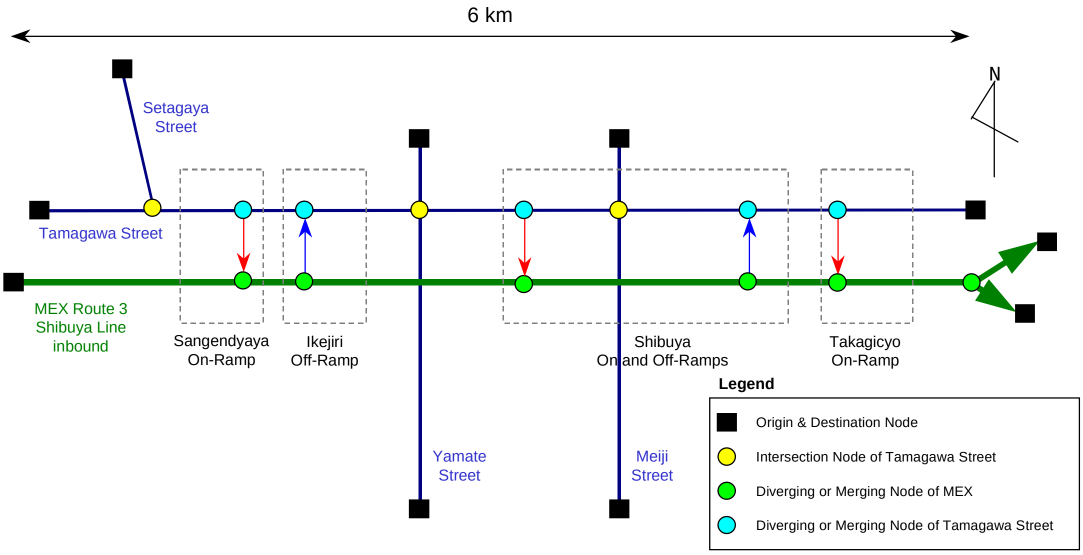

Figure 6 A Case Study Network

To validate the proposed ramp control strategy, we incorporate the strategy into AVENUE, a network traffic simulation model developed by our laboratory, and simulate by using the existing traffic conditions. Although we don't explain details of AVENUE here, it reproduces the accurate link density and the microscopic dynamic route choice behaviors of drivers at the same time. Figure 6 shows a case study network consisting of Tokyo Metropolitan Expressway (MEX) Route 3 Shibuya line inbound, Tamagawa Street, and so on. Results in Figure 7 prove the validity of the strategy.

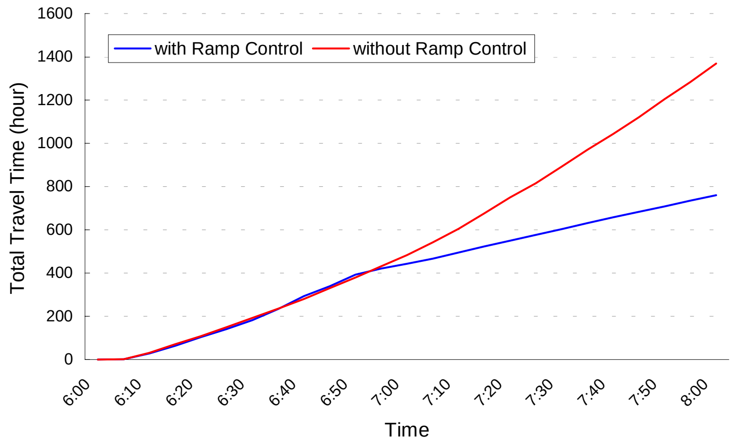

Figure 7 Comparison of Total Travel Times with and without Ramp Control

CONCLUSIONS

In this study, we propose the strategy for DSO considering the dynamic marginal cost and the optimum strategy of the ramp control toward DSO under the point queue concept. However, realistic physical queues cause some interesting phenomena that cannot be found with point queues. The ramp control analysis considering such a physical queue is our further work.