SYSTEM OPTIMIZING FLOW AND EXTERNALITIES IN TIME-DEPENDENT ROAD NETWORKS

Andy H.F. Chow

Centre for Transport Studies, University College London,
Gower Street, London WC1E 6BT, England.

ABSTRACT

This paper develops a framework for analysing and calculating system optimizing flow and externalities in time-dependent road networks. The externalities are derived by using a novel sensitivity analysis of traffic models. The optimal network flow is determined by solving a state-dependent optimal control problem, which assigns traffic such that the total system cost of the network system is minimized. This control theoretic formulation can work with general travel time models and cost functions. Deterministic queue is predominantly used in dynamic network models. The analysis in this paper is more general and is applied to calculate the system optimizing flow for Friesz's whole link traffic model. Numerical examples are provided for illustration and discussion. Finally, some concluding remarks are given.

1. INTRODUCTION

Dynamic traffic assignment models of route and departure time choice for travellers through congested networks provide important insight into the dynamics of peak periods and sensitivity of travellers' behaviour in response to a range of transport policy measures. In general, formulations of dynamic traffic assignment follow the extension of the two Wardrop's principles: user equilibrium and system optimum. The dynamic user equilibrium assignment has been the focus in the past two decades. As a result, we have already gained a substantial knowledge on the formulations, properties, and solution methods of dynamic equilibrium assignment.

This paper aims to analyse the dynamic system optimal assignment with departure time choice, which is an important, yet underdeveloped area. The dynamic system optimal assignment process suggests that there is a central "system manager" to distribute network traffic over time within a fixed horizon. Consequently, the total, rather than individual, travel cost of all travellers through the network is minimised. Although the system optimal assignment is not a realistic representation of network traffic, it provides a bound on how we can make the best use of the road system, and as such it is a useful benchmark for evaluating various transport policy measures.

Proceeding after Heydecker and Addison (2005), the travel cost incurred by each traveller is considered to have three distinct components: a cost related to the travel time *en route*, and time-specific costs associated with the departure time of the traveller from the origin and the arrival time at the destination respectively. Given the assigned network flow, the associated travel times through the network are determined by a traffic model. The travel times then influence the arrival times of travellers and hence the travel costs incurred. Following the Daganzo (1995) and Mun (2001), to ensure the satisfaction of several necessary physical principles such as proper flow propagation (or consistency between flows and travel times), non-negativity of flow, first-in-first-out (FIFO) queue discipline, and causality, the traffic model adopted in this paper considers the travel time on each link to be a linear non-decreasing function of link traffic volume.

Many previous analyses on dynamic system optimum and externalities adopted an optimal control theoretic formulation with Merchant and Nemhauser's (1978) outflow traffic model. On the one hand, this formulation provides some attractive mathematical properties for analysis. On the other hand, however, it ignores the importance of ensuring proper flow propagation. In addition, the outflow models have also been widely criticized for their implausible traffic behaviour (see Astarita, 1996; Heydecker and Addison, 1998; Mun, 2001).

In addition to the system optimizing flow, it is noted that each additional traveller, who enters the system at a certain time, imposes an additional travel cost on the others who enter the system at that time and thereafter. Understanding the nature of this externality is important in managing dynamic network. Previous research on the externality was specific to certain kinds of traffic models. Some traffic models adopted in some previous studies were even now considered to be implausible for various reasons. This paper revisits the dynamic externality in a more general and plausible way. We develop a novel sensitivity analysis of the traffic models and apply it to derive the externality in an optimal control theoretic formulation.

In Section 2, we first present the formulation and necessary conditions of the dynamic system optimal assignment in the next section. The dynamic system optimal assignment problem is formulated as a state-dependent optimal control problem. Following Friesz et al., (2001), we use this to analyse and solve dynamic system optimal assignment problem. To solve the dynamic system optimal flow and analyse the associated flow externalities, a novel sensitivity analysis of the traffic model with respect to the link inflow is adopted. The sensitivity analysis is developed through flow propagation mechanism and the analysis is not confined to a specific traffic model. Indeed, we apply the sensitivity analysis to deterministic queuing model and we are managed to restore previous analytical results (Ghali and Smith, 1993; Kuwahara, 2001). Then, solution algorithms are presented for implementing the sensitivity analysis and solving the dynamic traffic assignments. With the solution algorithms, we show some numerical calculations and discuss the characteristics of the results. Finally, some concluding remarks are given.

2. SYSTEM OPTIMAL ASSIGNMENT AND EXTERNALITY

2.1 Formulation

The system optimal assignment with departure time choice for fixed travel demand can be formulated as the following optimal control problem. The optimization problem (1) – (6) looks for an optimal inflow profile, $e_a(s)$, which minimizes the total system travel cost within the planning horizon given a fixed amount of total throughput:

$$\min_{e_a(s)} Z = \sum_{\forall a} \int_0^T C_a(s) e_a(s) ds \tag{1}$$

subject to:

$$\frac{dG_a[\tau_a(s)]}{ds} = e_a(s) \quad ,\forall s, \forall a \tag{2}$$

$$\frac{dx_a(s)}{ds} = e_a(s) - g_a(s) \quad ,\forall s, \forall a \tag{3}$$

$$\frac{dE_a(s)}{ds} = e_a(s) \quad ,\forall s, \forall a \tag{4}$$

$$\sum_{\forall a} E_a(T) = J_{od} \tag{5}$$

$$e_a(s) \geq 0 \quad ,\forall s, \forall a \tag{6}$$

The objective function was adopted by Merchant and Nemhauser’s (1978), and by several other researchers since then. Proceeding after Heydecker and Addison (2005), this study considers the total travel cost $C_a(s)$ encountered by each traveller on the travel link has three distinct components. The first component is the time spent on travelling along the link, which is determined by the travel time model that is adopted. In addition to the travel time, we add a time-specific cost $f[\tau_a(s)]$ associated with arrival time $\tau_a(s)$ through route p at the destination. Finally, we add a time-specific cost $h(s)$ associated with departure from the origin at time s. Possible choices of these time-specific cost functions are investigated by Heydecker and Addison (2005). Consequently, the total travel cost $C_a(s)$ associated with entry time to link a at time s is determined as a linear combination of these costs as

$$C_a(s) = h(s) + [\tau_a(s) - s] + f[\tau_a(s)]. \tag{7}$$

The notation $\tau_a(s)$ denotes the exit time from the link a for traffic which enters at time s. Following Daganzo (1995) and Mun (2001), we consider the exit time $\tau_a(s)$ to be a linear non-decreasing of link traffic volume $x_a(s)$, hence plausible traffic behaviours such as FIFO queue discipline and proper flow propagation are guaranteed. Following Friesz et al.'s (1993), we consider that $\tau_a(s)$ takes the functional form as

$$\tau_a(s) = s + \phi_a + \frac{x_a(s)}{Q_a}, \tag{8}$$

where the amount of whole link traffic at time s is represented by $x_a(s)$. The free flow travel time and the capacity of the travel link are denoted by ϕ_a and Q_a respectively. Equations (2) ensure the proper flow propagation along each route, in which $G_a(s)$ denotes the cumulative outflow by the exit time $\tau_a(s)$. Equations (3) are the state equations that govern the evolution of link traffic, $x_a(s)$. The variables $e_a(s)$ and $g_a(s)$ represent the flow rates at time s of inflow and outflow respectively. Equations (4) define the cumulative inflow $E_a(s)$. Equation (5) specifies the amount of total throughput J_{od} generated in the system within the time horizon T. Conditions (6) ensure the positivity of the control variable, $e_a(s)$, for all times s. Given a positive inflow $e_a(s)$, the corresponding outflow $g_a(s)$ and link traffic volume $x_a(s)$ is guaranteed to be positive (see for example, Mun, 2001). Hence, we do not add explicit constraints to ensure the non-negativity of $g_a(s)$ and $x_a(s)$. The class of traffic models considered in this paper has been shown to satisfy FIFO structurally (see, for example, Daganzo, 1995; Mun, 2001), we do not need to add any explicit constraint for this.

2.2 Analysis

One technical difficulty is that with the traffic models above, the time lag between changes to the control variable, $e_a(s)$, and the corresponding responses, $g_a(s)$, is state-dependent. This state-dependent control theoretic formulation is unorthodox. Its properties and application to dynamic equilibrium were studied by Friesz et al. (2001). As an extension to Friesz et al. (2001), the necessary conditions for the state-dependent system optimization problem are given by the following the proposition.

Proposition 1: The necessary conditions for the optimization problem (1) – (6) can be derived as

$$e_a(s)\begin{cases} > 0 \Rightarrow C_a(s) + \int_0^T \left.\frac{\partial C_a}{\partial u_s}\right|_t e_a(t)dt + \lambda_a(s) - \lambda_a[\tau_a(s)] = \mu_a(s) = \nu_{od} \\ = 0 \Rightarrow C_a(s) + \int_0^T \left.\frac{\partial C_a}{\partial u_s}\right|_t e_a(t)dt + \lambda_a(s) - \lambda_a[\tau_a(s)] \geq \mu_a(s) = \nu_{od} \end{cases} ,\forall a, \forall s \in [0,T] \tag{9}$$

where $\mu_a(s) = \nu_{od}$ is constant with respect to time and its magnitude is determined by the predefined amount of throughput.

Proof:
See Appendix A in Chow (2006). ☐

The notation $\lambda_a(s)$ and $\lambda_a[\tau_a(s)]$ denote the costate variables at times s and $\tau_a(s)$ respectively, where

$$\lambda_a(s) = \frac{1}{Q_a}\int_{t=s}^{T}(1+f'[\tau_a(t)])e_a(t)dt\,. \tag{10}$$

This costate variable represents the sensitivity of the value of the objective function with respect to the changes in state variable $x_a(\cdot)$ at the associated time.

The first term on the left-hand-side of (9), $C_a(s)$, is the cost experienced by that additional traveller given the current traffic condition, and the integral in the second term on the left-hand-side of (9), $\Psi_a(s) = \int_0^T \left.\frac{\partial C_a}{\partial u_s}\right|_t e_a(t)dt$, is the additional travel cost, which is regarded as externality, imposed by an additional amount of traffic, u_s, at time s to existing travellers in the system. Capturing the externality is important in managing dynamic network, and it requires knowing the sensitivity of the total travel cost $\left.\frac{\partial C_a}{\partial u_s}\right|_t$ for each departure time s with respect to a change of u_s in the link inflow a particular time s. In this study, we consider parameters u_s of the form for which

$$\frac{de_a(t)}{du_s} = \begin{cases} 1 & \text{if } t \in [s, s+ds) \\ 0 & \text{otherwise} \end{cases}, \tag{11}$$

in which ds represents the incremental time step. Differentiating both sides of (7) with respect to u_s, we have

$$\left.\frac{\partial C_a}{\partial u_s}\right|_t = (1+f'[\tau_a(t)])\left.\frac{\partial \tau_a}{\partial u_s}\right|_t\,. \tag{12}$$

3. SENSITIVITY ANALYSIS OF TRAFFIC MODELS

Calculating the externality $\Psi_a(s)$ requires the sensitivity of traffic models with respect to perturbations in link traffic inflow. Consequently, the section derives a novel expression for the sensitivity of the time of exit with respect to such perturbations in inflow, which is given in the following proposition.

Proposition 2: Suppose there is a change of u_s in the link inflow rate at a particular time s, the sensitivity of the time of exit at a time s with respect to this perturbation can be calculated as

$$\left.\frac{\partial \tau_a}{\partial u_s}\right|_t = \frac{1}{Q_a}\left\{\int_{\kappa=\sigma_a(t)}^{t}\frac{de_a(\kappa)}{du_s}d\kappa + g_a(t)\left.\frac{\partial \tau_a}{\partial u_s}\right|_{\sigma_a(t)}\right\}. \tag{13}$$

Proof:
See Appendix B in Chow (2006). ☐

The derivative of exit time with respect to the change of u in inflow is then expressed in terms of the dependence of the inflow profile $e_a(s)$ in which s lies between s and $\sigma_a(s)$, the current outflow $g_a(s)$, and the derivative of exit time at the time of entry, $\sigma_a(s)$.

4. SOLUTION ALGORITHM

We propose the following procedure to solve for the dynamic system optimal assignment with fixed travel demand:

Step 0: Initialisation

0.1 Choose an initial equilibrium cost C^*_{od};

0.2 Set the overall iteration counter $n := 1$;

0.3 Set $e_a(k) := 0$ for all links a, $a \in [1, A]$, and all times k, $k \in [0, K]$. The notation $e_a(k)$ represents the assigned inflow to link a between times $k\Delta s$ and $(k+1)\Delta s$. The total number of simulated time steps is denoted as $K = T/\Delta s$ and the total number of parallel links is denoted by A; set time index $k := 0$;

0.4 Set costates $\lambda_a(k) := 0$ for all times $k \in [0, K]$;

0.5 Set the link index $a := 1$;

0.6 Set the time index $k := 0$;

0.7 Set the overall iteration counter $n^i := 1$.

Step 1: Network loading

Find $\tau_a(k+1)$ by loading the travel link using the inflow $e_a(k)$ at the current iteration. The network loading algorithm "Algorithm D2" described in Nie and Zhang (2005) was adopted for this purpose.

Step 2: Calculating externality

Use Algorithm 2 to calculate the externality $\Psi_a(k)$ associated with each $e_a(k)$.

Step 3: Determining the auxiliary inflow

3.1 Calculate

$$C_a(k+1) = h(k+1) + [\tau_a(k+1) - (k+1)] + f[\tau_a(k+1)] + \Psi_a(k+1) + \lambda_a(k) - \lambda_a[\tau_a(k)];$$

3.2 Calculate $\Omega = \dfrac{C_a(k+1) - C_a(k)}{\Delta s}$ and $\Omega' = \dfrac{\partial \Omega}{\partial e_a(k)} = (1 + f'[\tau_a(k+1)])\dfrac{1}{Q_a}$,

in which $f'[\tau_a(k)] \approx \dfrac{f[\tau_a(k+1)] - f[\tau_a(k)]}{\tau_a(k+1) - \tau_a(k)}$ using a finite difference approximation;

3.3 Calculate the auxiliary inflow, $d_a(k) = -{}^{\Omega_a(k)}\!/\!_{\Omega'_a(k)}$, with the second-order searching direction.

Step 4: Stopping criteria for calculating auxiliary flow

4.1. Check if $\left|C_a(k+1) - C^*_{od}\right| \le \varepsilon$ or n^i is greater than the predefined maximum number of inner iterations, then go to step 4.2; otherwise, set $n^i := n^i + 1$ and go to step 1.

4.2. If $k = K$, then go to step 4.3; otherwise $k := k + 1$ and go to step 1;

4.3. If $a = A$, then go to step 4.4; otherwise $a := a + 1$ and go to step 0.6;

4.4. Check if the total throughput $E_{od} = \sum_{\forall a}\sum_{\forall k} e_a(k)$ from the system is equal to the predefined total demand J_{od} for the *o-d* pair. If yes, then *terminate* the algorithm; otherwise update $C_{od}^* := C_{od}^* + \left[\dfrac{J_{od} - E_{od}}{\dfrac{dE_{od}}{dC^*}}\right]$, and go back to step 0.3. The derivative $\dfrac{dE_{od}}{dC_{od}^*}$ is given by Heydecker (2002).

Step 5: Determining step size

Search for an optimal step size θ by using golden section method and update the inflow as $e_a(k) := \max[e_a(k) + \theta d_a(k), 0]$ for all times *s* such that the total system cost is minimized.

Step 6: Calculating the associated costate variables

6.1 Set the link index $a := 1$;

6.2 Set $\lambda_a(K) = 0$;

6.3 Set the time index $k := K - 1$;

6.4 Compute $\lambda_a(k) = \lambda_a(k+1) + (1 + f'[\tau_a(k)])\dfrac{e_a(k)}{Q_a}\Delta s$;

6.5 Calculate $\lambda_a[\tau_a(k)]$ from $\lambda_a(k)$ and $\tau_a(k)$ using linear interpolation as

$$\lambda_a[\tau_a(k)] \approx \lambda_a\big|_{\lfloor \tau_a(k) \rfloor} + \left(\lambda_a\big|_{\lceil \tau_a(k) \rceil} - \lambda_a\big|_{\lfloor \tau_a(k) \rfloor}\right)\left(\tau_a(k) - \lfloor \tau_a(k) \rfloor\right).$$

6.6. If $k = 0$, then go to step 6.7; otherwise $k := k - 1$ and go to step 6.2;

6.7. If $a = A$, then go to step 7; otherwise $a := a + 1$ and go to step 6.1.

Step 7: Overall stopping criteria

Define $\xi = \dfrac{\sum_{\forall a}\sum_{\forall k} e_a(k)\left|C_a(k+1) - C_{od}^*\right|}{\sum_{\forall a}\sum_{\forall k} e_a(k) C_{od}^*}$ as a measure of disequilibrium, which is equal to zero at system optimum. If n is greater than the predefined maximum number of overall iterations or ξ is sufficiently small, i.e. $\xi \le \varepsilon$ where ε is a test value, then go to stop; otherwise set $n := n+1$ and go to step 1.2.

5. NUMERICAL EXAMPLES

We first consider a single link, which has a free flow time 3 mins and a capacity 20 vehs/min, connecting a single origin-destination pair. The size of discretized time interval Δs is taken as

1 min. We first show the numerical solutions of the whole link traffic model. A parabolic profile, which is specified as (14), of inflow is loaded into the travel link.

$$e_a(s) = \begin{cases} \frac{1}{8}(40-s)s & \text{if } 0 \le s \le 40 \\ 0 & \text{otherwise} \end{cases} . \tag{14}$$

To investigate the accuracy of the sensitivity analysis, the parabolic inflow profile is perturbed at time 1. The associated variations in travel time are plotted in Figure 1. The variations are calculated according to equation (13). In the same figure, the variations determined by using numerical finite difference method are also plotted for comparison. To calculate the finite difference, one extra unit of inflow is added at time 1, others remain unchanged. The variations in travel times are then calculated by repeated link loading with the original inflow profile versus the perturbed inflow profile. The result shows that the analytical variations given by equation (13) can represent the true variations in travel time reasonably well. Both numerical and analytical derivatives drop to zero at time 83 when all traffic is cleared from the link.

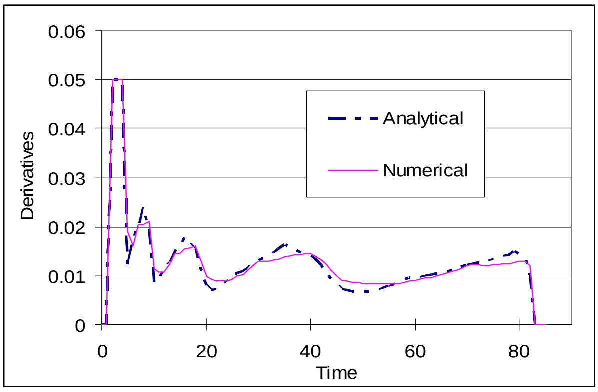

Figure 1 Sensitivity of travel time with respect to a perturbation in inflow

To calculate the dynamic traffic assignments, one more travel link having a free flow travel time 4 mins and a link capacity 30 vehs/min is added to the system. Furthermore, the origin-specific cost is considered to be a monotone linear function of time with a slope -0.4. The destination cost function is piecewise linear, with no penalty for arrivals before the preferred arrival time $t^* = 50$, and increases with a rate 2 afterwards. The length of the planning horizon $[0, T]$, where T=100, is set such that that all traffic can be cleared by time T. The total amount of traffic J_{od} is taken as 800 vehs. Figure 2 shows the corresponding profiles of inflow and outflow, and the total travel cost at equilibrium. The traffic is assigned to the route 1 during times 18 and 49, and to route 2 during times 21 and 49. The route flow volumes using route 1 and route 2 are 380.25 (vehs) and 419.75 (vehs) respectively. The measure of disequilibrium ξ achieved is below 10^{-17}. At dynamic user equilibrium, the total system travel cost is 12,465.2 veh-hr.

Figure 3 shows the assignment of the dynamic system optimum. With the same total throughput J_{od}, the period of assignment to route 1 expands from times [18, 49] to times [4, 56], while that to route 2 expands from times [21, 49] to times [6, 50]. In general, the profiles of route inflows are more spread at system optimum to reduce the intensity of congestion on the routes, whilst maintaining the same volume of travel. The associated total system travel cost at system optimum is decreased from 12,465.2 veh-hr in user equilibrium to 11,447.3 veh-hr. However, due to the addition of externality and the costate variables, the marginal social cost at which travel takes place increases from 15.58 min at user equilibrium to 21.78 min at system optimum, although the system optimizing flow causes the decrease in total system cost.

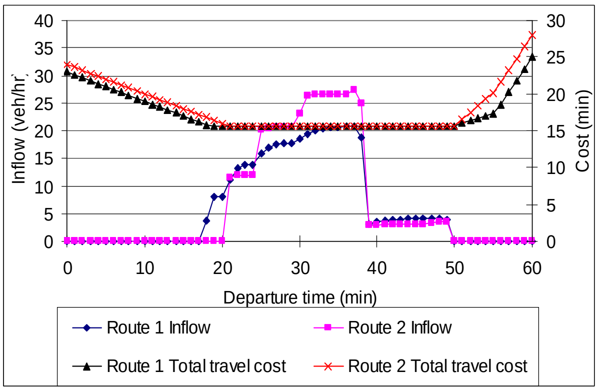

Figure 2 Equilibrium assignment

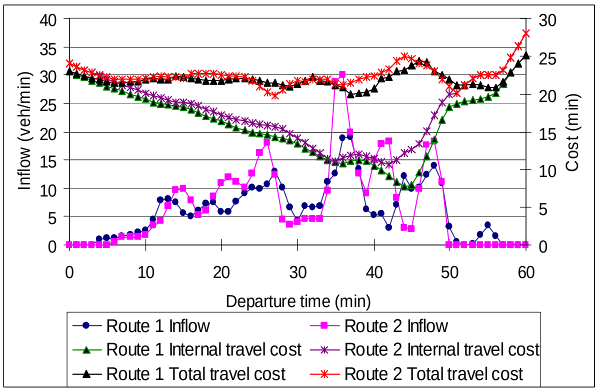

Figure 3 System optimal assignment

To illustrate the cause of the decrease in system travel cost, Figure 4 shows the link traffic volumes at user equilibrium and at system optimum. Furthermore, using the sensitivity analysis of traffic models described in proposition 2, the externalities imposed by travellers to the system at equilibrium condition and system optimal condition are also calculated and the numerical results are plotted in Figure 5. Interestingly, yet importantly, the results show that, with Friesz et al's (1993) travel time model, the system optimal assignment has to allow queuing, and the externality that each traveller imposes on the others is not zero even at system optimum. The system optimal assignment can only manage and minimize queuing and externality of each traveller imposes on the others. This implies that the previous analyses on dynamic system optimum using the deterministic queuing model do not apply in general.

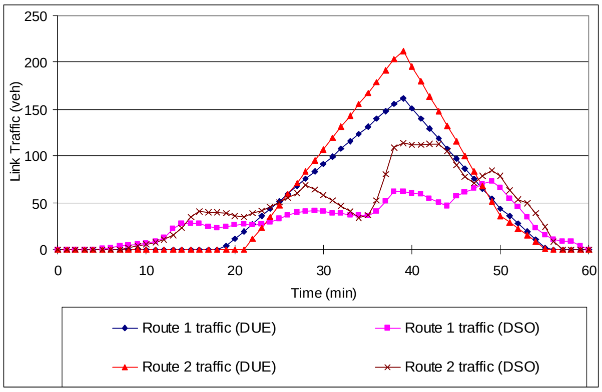

Figure 4 Link traffic volumes

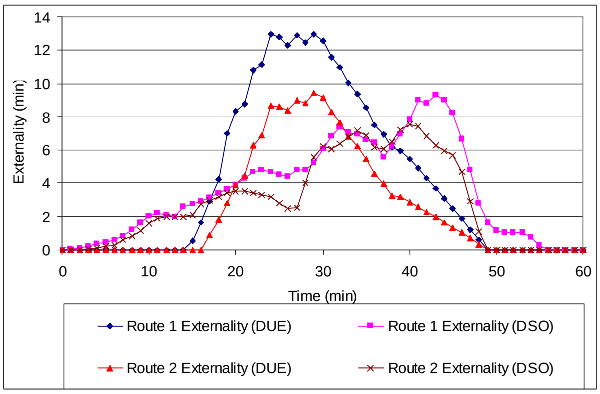

Figure 5 Externalities

6. CONCLUDING REMARKS

The main contribution of this paper is the necessary conditions for dynamic system optimizing flow. To solve the system optimal assignment, we also developed a novel sensitivity analysis of traffic models with respect to perturbations in link inflow. We then presented solution algorithms for implementing the sensitivity analysis and solving the dynamic traffic assignments. We also applied the algorithms to numerical calculations. The characteristics of the results were discussed. In particular, with Friesz et al's (1993) linear traffic model, the system optimal assignment has to allow queuing, and the externality that each traveller imposes on the others is not zero even at system optimum. We can only manage and minimize queuing and externality of each traveller imposes on the others. This implies that the previous analyses on dynamic system optimum using the deterministic queuing model do not apply generally.

The study gives us a deeper understanding of the nature of dynamic system optimal assignment on a plausible framework. In addition to the system optimizing flow, this paper revisited the dynamic externality in a more general and plausible way. Furthermore, in the present study, the formulation and analysis presented are restricted to networks in which capacity limitations of different routes are mutually distinct. We are currently exploring ways in which this analysis can be extended to consider shared bottlenecks in general networks.

ACKNOWLEDGEMENTS

The continuing guidance and stimulating discussions from Professor Benjamin Heydecker and Dr. JD Addison are gratefully appreciated.

REFERENCES

Astarita, A (1996) A continuous time link model for dynamic network loading based on travel time functions. J-B Lesort, ed. *Transportation and Traffic Theory*. Pergamon, Oxford, 79-102.

Chow, AHF (2006) Analysis of dynamic system optimum and externalities with departure time choice. *Working paper,* University College London (ISSN: 1747-6232).

Daganzo, CF (1995) Properties of link travel time functions under dynamic loads. *Transportation Research Part B*, **29**(2), 95-98.

Friesz, TL, Bernstein, D, Smith, TE, Tobin, RL and Wie, BW (1993) A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research*, **41**(1), 179-91.

Friesz, TL, Bernstein, D, Suo, Z and Tobin, RL (2001) Dynamic network user equilibrium with state-dependent time lags. *Networks and Spatial Economics* **1**(3/4), 319-47.

Heydecker, BG (2002) Dynamic equilibrium network design. MAP Taylor, ed. *Transportation and Traffic Theory*. Pergamon, Oxford, 349-70.

Heydecker, BG and Addison, JD (1998) Analysis of traffic models for dynamic equilibrium traffic assignment. MGH Bell, ed. *Transportation Networks: Recent methodological Advance*. Pergamon, Oxford, 35-49.

Heydecker, BG and Addison, JD (2005) Analysis of dynamic traffic equilibrium with departure time choice. *Transportation Science,* **39**(1), 39–57.

Merchant, DK and Nemhauser, GL (1978a) A model and an algorithm for the dynamic traffic assignment problem. *Transportation Science,***12**(3), 183-199.

Mun, JS (2001) A divided linear travel time model for dynamic traffic assignment. *9th World Conference on Transport Research*, Seoul, S.Korea.