



HAL
open science

DYNAMIC SYSTEM OPTIMAL ROUTING IN MULTIMODAL TRANSIT NETWORK

Tai-Yu Ma, Jean-Patrick Lebacque

► **To cite this version:**

Tai-Yu Ma, Jean-Patrick Lebacque. DYNAMIC SYSTEM OPTIMAL ROUTING IN MULTIMODAL TRANSIT NETWORK. 2012. hal-00740347

HAL Id: hal-00740347

<https://hal.science/hal-00740347>

Preprint submitted on 9 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DYNAMIC SYSTEM OPTIMAL ROUTING IN MULTIMODAL TRANSIT NETWORK

Tai-Yu Ma*

Laboratoire d'Economie des Transports (LET), CNRS-Université de Lyon

Fax (33) 4-72-72-64-48, Phone (33) 4-72-72-64-46

E-mail: tai-yu.ma@let.ish-lyon.cnrs.fr

14, Avenue Berthelot

F-69363 Lyon, France

Jean-Patrick Lebacque

UPE (Université Paris Est)

IFSTTAR Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux (French institute of science and technology for transport, development and networks)

GRETTIA Génie des Réseaux de Transports Terrestres et Informatique Avancée (Engineering of Surface Transportation Networks and Advanced Computing)

Fax (33) 1-45-92-55-00, Phone (33) 01-45-92-56-26

Email : jean-patrick.lebacque@ifsttar.fr

Le Descartes 2, 2 rue de la Butte Verte

93166 Noisy-le-Grand Cedex, France

*Corresponding author

Number of words in text: 5255 words

Number of figures and tables: 7 (1750 words)

Total: 7005 words

ABSTRACT

The system optimal routing problem has been widely studied for road network while it is less considered for public transit system. Traditional shortest-path-based multimodal itinerary guidance systems may deteriorate the system performance when the assigned lines become congested. For this issue, we formulate the dynamic system optimal routing model for multimodal transit system. The transit system is represented by a multilevel graph to explicitly simulate passenger flow and transit system operations. A solution algorithm based on the cross entropy method is proposed, and its performance is compared with the method of successive averages in static and dynamic cases. Numerical study on a simple multimodal transit network provides the basis for comparing the system optimal routing and user optimal routing under different congestion levels.

INTRODUCTION

The development of internet enabled service and information system for assisting travellers to reach his destination in a multimodal transit network has been an active research issue in recent years. The equipped travelers are provided with multimodal itinerary recommendations during his/her journey to destination based on the timetable/frequency or real-time information. More advanced services concern on-line ticket booking and delivery on smart phone to enhance a more efficient passengers' transportation in the multimodal transit system. As there is an increasing demand in congested urban area, such system can effectively reduce users' travel time and improves the efficiency of multimodal transit system.

Existing multimodal itinerary guidance systems can be found in (1, 2, 3). These systems provide passengers with pre-trip multimodal itineraries by searching time-dependent shortest paths to requested destinations. Different multimodal network models have been proposed for computing the shortest multimodal paths. For example, Foo et al. (1) developed a multicriteria and multimodal pre-trip advisory system based on the computation of the shortest paths in terms of generalized path travel cost. The multimodal transport system is modeled on a directed graph for which each service line section is represented by a direct arc connecting two service stops. Lo et al. (2) developed a multimodal route advisory system based on the State-Augmented Multimodal (SAM) network with frequency-based transit operations. The SAM network represents modal transfer states of path by some internal state labels associated with each node, which is convenient for restricting mode transfer constraints with non-linear path fare. Galvez-Fernandez et al. (3) proposed a transfer graph consisting of a set of unimodal graphs connected by a set of transfer links. The time-dependent multimodal shortest path advice is computed based on the variant of Dijkstra's algorithm by linking the shortest unimodal subpaths via related transfer points. Although these multimodal transit itinerary advice systems utilized different network loading models for representing passengers' flow and transit vehicle operations, the route guidance principle is based on the time-dependent shortest path assignment. This assignment scheme may increase total system travel time if the informed users exceed certain proportion and reduce the system performance.

For this issue, previous studies (4) on road route guidance systems suggest that the shortest path based route guidance is effective in reducing user's travel time for few equipped vehicles but not for a large proportion of equipped vehicles. As the transit vehicle has limited passenger capacity, the assignment of too many users on the same shortest path will lead to undesired congestions and increase total system travel times. Effective transit route guidance system should take into account this effect. For this issue, Jahn et al. (5) propose a system-optimal (SO) routing with user-constrained shortest paths on static road network. They considered the reactive route guidance problem as a convex minimum cost

multicommodity problem with side constraints. The assigned routes are restricted to a set of paths with a reasonable higher travel time than that on the shortest paths. The computational study indicated that such route guidance system provides significant advantage in fairness (less travel time deviation between users) and in efficiency (total travel time of users is close to the system optimum). For dynamic road route guidance, some more advanced approaches such as anticipatory route guidance can be applied by taking into account future traffic state predictions under the according route guidance (6). Although the dynamic route guidance problem has been widely studied for road users, it is still less considered in existing transit route advisory systems.

In this study, we propose a dynamic system optimal routing (DSOR) on multimodal transit network. We assume that all passengers are equipped with some communication advices and follow the suggested pre-trip multimodal itinerary to his/her related destination. Although this assumption is relatively strong, it can be considered as the reference scenario for the design of dynamic user-constrained or online consistent multimodal itinerary advice systems. The system optimal routing problem under full compliance and full market penetration is equivalent to the system optimal dynamic traffic assignment problem (SODTA). The SODTA problem aims at determining time-dependent path flows such that total system travel time/cost is minimized.

In the past, the SODTA problem on road network has been widely studied (6, 7, 8). Beckman et al. (9) first introduce the concept of *path marginal cost* (time) to describe the system optimal state. The path marginal cost is the induced extra travel cost on the system when introducing a user/vehicle on a path. Under system optimal state, the experienced path marginal cost is equal and no more than that of unused paths. As there is no direct way to compute the path marginal cost under simulation-based dynamic network loading models, it is generally difficult to evaluate its value. Peeta and Mahmassani (10) show that this problem can be formulated as a path-based dynamic user equilibrium (DUE) assignment problem and solved by related solution methods for the DUE problem. The authors compute the link marginal cost by approximating the derivative of link travel times. Chow (7) proposes a system optimal traffic assignment model with departure time choice, and formulates the problem as a state-dependent optimal control problem under different link travel time models. The author proposes a hybrid gradient-based forward–backward dynamic programming approach for solving the problem in a small network. The widely used solution method, among many others (see (8) for the literature review), for solving the dynamic SO assignment problems is the method of successive averages (MSA) (11, 12). It is well known that the MSA method can obtain the approximate of DUE solution for simulation-based network loading model. To the best of our knowledge, there is still no related study on dynamic system optimal assignment on the multimodal transit network.

The rest of this paper is organized as follows. First the dynamic network loading

model based on a multilevel multimodal network is presented. Then we illustrate the computation of generalized path cost and the multiagent-based transit system simulation. It follows the mathematical formulation of the DSOR problem on the multimodal transit system. The computation of time-dependent path marginal cost on the multimodal transit system is discussed. In section 4, the solution algorithm based on the cross entropy (CE) method is presented (13, 14, 15). In Section 5, we validate the proposed solution method on a static small network and compare its solution quality and convergence speed with the MSA method. Then we applied the solution method for the DSOR problem and compare its performance with the MSA method under different travel demand settings. Finally, conclusions are drawn and future extensions are discussed.

DYNAMIC MULTIMODAL TRANSIT NETWORK LOADING MODEL

Multimodal Transit Network

The multimodal transit network is represented by a direct graph $G(N, A)$ with N the set of nodes and A the set of arcs (Figure 1). The multimodal network is modelled as a multilevel structure where each level represents a unimodal subnetwork and connected between them by a set of transfer links (16, 17). The reference level is the origin/destination network where each node represents an origin and/or destination connected between them via walking links. Each transit mode operates on its own subnetwork for which we distinguish *station node* and *line node (stop)* for modelling passengers' and vehicles' flows. The station nodes are mode-specific and interconnected by boarding /alighting arcs with related service line nodes. Moreover, within a multimodal station, a set of station nodes are used to connect related service lines, and they are also interconnected by transfer links. The walking and transfer links are associated with constant access travel time. It is calculated as its length divided by constant walking speed. As for the transit lines, they are a set of pre-defined sequence of line nodes for transit vehicle operations. The travel time on transit line arcs is calculated by its length divided by constant average vehicle operation speed. Hence, a multimodal path is explicitly represented by acyclic directed path on the multilevel directed graph connecting an OD pair. The transit congestion is considered under the constraints of vehicle capacity.

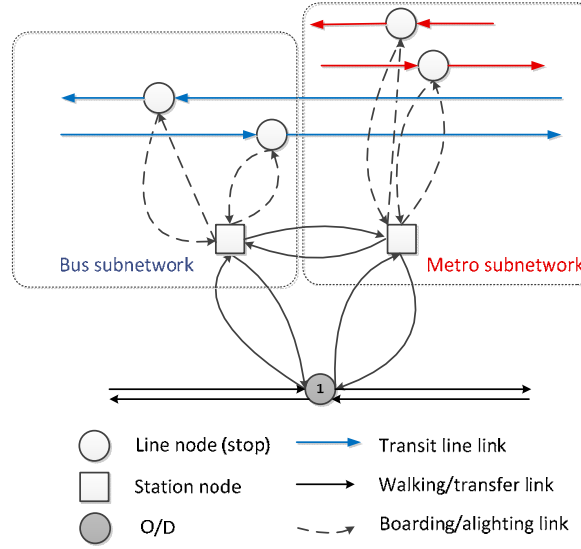


FIGURE 1 Multilevel transit network structure

Dynamic Network Loading and the Computation of the Generalized Cost of Multimodal Transit Path

For modelling the transit system operation dynamics, we adopt a multiagent approach to capture the dynamics of transit system operations and passenger's waiting process at stations. Two types of agents are used for modeling passengers and transit vehicle movements. For the vehicle agents, it represents a mode-specific vehicle operating on predefined transit lines under vehicle capacity constraints and scheduled service frequency. The stop times at stations are assumed constant and sufficient for passengers to board or alight the vehicle. If the vehicle capacity is achieved, the passengers have to wait for the next arriving vehicle. The passenger agents aim at arriving at his/her destination with the least travel cost following an acyclic path in the multimodal transit network. In general, passenger's experienced path travel cost contains total path travel time, waiting time at stops, transfer penalty, schedule delay cost of early/late arrival at destination and fare. For simplicity, the last three terms of the generalized travel cost is neglected. Hence, the generalized travel cost of a multimodal path contains the walking time for accessing O/D and stations, transfer time between stations, boarding/alighting time, and in-vehicle travel time and waiting time at line (stop) nodes. By assuming the First-In-First-Out principle for passengers' queuing process at stops, the waiting times $\pi_i^w(t)$ when arriving at a stop i at time t can be calculated by:

$$\pi_i^w(t) = D_i^{-1}(S_i(t)) - t, \quad (1)$$

where $S_i(t)$ is cumulative arrivals at line node i by time t , $D_i^{-1}(t)$ is the inverse function of

cumulative departure from line node i by time t .

Hence the generalized travel cost of path r with respect to departure time interval h and OD pair k is then evaluated as:

$$C_{hkr}(\mathbf{f}) = \frac{1}{|M_{hkr}|} \sum_{m \in M_{hkr}} C_m(\mathbf{f}) = \frac{1}{|M_{hkr}|} \sum_{m \in M_{hkr}} [\pi_r^s + \pi_r^v + \pi_r^w(t_m^{\text{dep}}, \mathbf{f})], \quad (2)$$

where $C_m(\mathbf{f})$ is passenger m 's experienced travel times. π_r^s is the total walking times of path r , π_r^v the in-vehicle travel times on transit links of path r , and $\pi_r^w(t_m^{\text{dep}}, \mathbf{f})$ the total experienced waiting times on path r , calculated as the summation of the waiting times at line nodes of path r when agent m departing from his/her origin at time t_m^{dep} . Note that the waiting times at stops depend on related line frequency, vehicle capacity and the implying queuing process at stops.

PROBLEM FORMULATION

Notation

m	designation of a user
h	departure time index with discretized time slot Δ , $h \in H = \{0, 1, \dots, n\}$ where $\lceil T_h / \Delta \rceil = n$ with T_h being the total period for departure time choice
k	origin-destination pair, $k \in K$
r	path index $r \in R$
d_{hk}	demand of origin-destination pair k in departure time interval h
$C_{hkr}(\mathbf{f})$	experienced path generalized travel cost with respect to departure time interval h , OD pair k and path r
f_{hkr}	flow on path r connecting OD pair k in departure time period h
t	time instant
T	the time of the last vehicle/user leaves the network

We consider the DSOR problem with given time-dependent OD demand $d_{hk}, \forall h, k$. The problem aims at determining time-dependent path flow patterns such that total system travel cost (time) is minimized. The problem can be formulated as the following minimization problem (Peeta and Mahmassani (10)):

$$\text{Min } Z(\mathbf{f}) = \sum_h \sum_k \sum_{r \in R_{hk}} f_{hkr} C_{hkr}(\mathbf{f}) \quad (3)$$

$$\text{s.t. } \sum_{r \in R_{hk}} f_{hkr} = d_{hk}, \forall h \in H, k \in K \quad (4)$$

$$f_{hkr} \geq 0, \forall h \in H, k \in K, r \in R_{hk} \quad (5)$$

The objective function (3) represents the minimization of total generalized travel cost of users, calculated as the summation of the multiplication of the path flows and its average time-dependent generalized travel cost over all OD pairs and the study period. The constraint (4) is the conservation of path flows for all OD pairs and departure time intervals given the known time-dependent OD demand. The constraint (5) ensures the non-negativity of path flow. As the generalized travel cost depends on complex interactions of transit system supply and users' demand, it is difficult to obtain an analytical form of travel cost estimation. The experienced travel cost is generally calculated by a simulator which describes the dynamics of transit vehicles and users on the multimodal network. Hence, the DSOR problem can not be solved directly by derivative-based optimization methods.

Alternatively, we can reformulate the original problem (3)-(5) by Lagrangian relaxation method as (10):

$$L(\mathbf{f}, \mathbf{u}) = \sum_h \sum_k \sum_{r \in R_{hk}} f_{hkr} C_{hkr}(\mathbf{f}) + \sum_h \sum_k u_{hk} (d_{hk} - \sum_{r \in R_{hk}} f_{hkr}) \quad (6)$$

$$\text{s.t.} \quad f_{hkr} \geq 0, \forall h \in H, k \in K, r \in R_{hk}, \quad (7)$$

where u_{hk} is the Lagrangian multiplier.

By taking the partial derivative of $L(\mathbf{f}, \mathbf{u})$ with respect to f_{hkr} , we obtain

$$\frac{\partial L(\mathbf{f}, \mathbf{u})}{\partial f_{hkr}} = C_{hkr}(\mathbf{f}) + \sum_{h'|h' \geq h} \sum_{k'} \sum_{r'} [f_{h'k'r'} \frac{\partial C_{h'k'r'}(\mathbf{f})}{\partial f_{hkr}}] - u_{hk} = \Gamma_{hkr} - u_{hk} \quad (8)$$

The first term Γ_{hkr} is the time-dependent path marginal cost with respect to h, k, r . It contains the generalized path cost $C_{hkr}(\mathbf{f})$ and the total marginal cost increases on the system from the moment when entering an additional user into the system. The second term u_{hk} is the dual variable representing the minimum path marginal cost with respect to h and k .

By introducing Eq. (8), the first-order optimality conditions are written as:

$$f_{hkr} (\Gamma_{hkr} - u_{hk}) = 0, \forall h, k, r \quad (9)$$

$$\Gamma_{hkr} - u_{hk} \geq 0, \forall h, k, r \quad (10)$$

Eq. (4) and (5)

Equations (9) and (10) state that, at system optimal state, the time-dependent marginal cost on used paths is less or equal than that on unused paths. Eq. (4) and (5) are the conservation of flows for each OD pair and the non-negativity of path flows, respectively. The solution for the above conditions is equivalent to that of the following variational inequality problem for

dynamic user equilibrium in terms of time-dependent path marginal cost $\Gamma(\mathbf{f}^*)$, i.e. find $\mathbf{f}^* \in \Omega$ such that

$$\Gamma(\mathbf{f}^*) \bullet (\mathbf{f} - \mathbf{f}^*) \geq 0, \forall \mathbf{f} \in \Omega \quad (11)$$

$$\Omega = \left\{ f_{hkr} \left| \sum_{r \in R_{hk}} f_{hkr} = d_{hk}, \forall f_{hkr} \geq 0, \forall h, k, r \right. \right\} \quad (12)$$

where the operator \bullet denotes the inner product of vectors.

The above reformulation of the DSOR problem as the variation inequality problem allows for the development of solution algorithms similar to that for solving dynamic user equilibrium problems. The difficulty remains on the issue of estimating the time-dependent path marginal cost on the system. We address this issue in the next section.

Computation of the Time-Dependent Marginal Cost of Multimodal Transit Paths

The time-dependent path marginal cost represents the increase of experienced generalized travel cost of the system when an additional passenger Δf_{hkr} is introduced on the path $r \in R_{hk}$ in departure time interval h . As passenger's generalized travel cost depends on the simulation-based dynamic network loading model, we cannot calculate directly the derivative of the generalized travel cost. As shown in (8), the time-dependent path marginal cost Γ_{hkr} is written as:

$$C_{hkr}(\mathbf{f}) + \sum_{h|h \geq h} \sum_{k'} \sum_{r' \in R_{h'k'}} [f_{h'k'r'} \frac{\partial C_{h'k'r'}(\mathbf{f})}{\partial f_{hkr}}] \quad (13)$$

The first term can be obtained as passengers' experienced average path generalized cost with respect to (h, k, r) . The second term can be estimated by first computing the time-dependent link marginal travel cost and then summing it over the path.

As shown in (2), passenger's experienced path cost depends on constant path walking times, path in-vehicle travel times and path waiting times at stops which depend on the available capacity when the passenger arriving at stops. If the number of passengers waiting at a stop exceeds the available capacity of the arriving vehicle, some passengers will be delayed and need to take the next arriving vehicle. The delay times equal to the additional waiting times due to failing to board the arriving vehicles. Hence we can write the approximate $\tilde{\Gamma}_{hkr}$ of (13) at each iteration as:

$$\tilde{\Gamma}_{hkr} = \frac{1}{f_{hkr}} \left[\sum_{m \in M_{hkr}} C_m(\mathbf{f}) + \sum_{i \in N_r^L} \Psi_h(Q_i(t)) \right], \quad (14)$$

where $\Psi_h(Q_i(t))$ is the total additional waiting times at stop i due to unavailable places of vehicle for passengers with departure time within h , i.e. $t^{\text{dep}} \in [t_0 + h\Delta, t_0 + (h+1)\Delta)$. N_r^L is the set of line nodes on path r . $Q_i(t)$ is the number of available places at line node i at time t . $\Psi_h(Q_i(t))$ can be computed by

$$\Psi_h(Q_i(t)) = \sum_k \sum_{r \in R_{nk}} \sum_{m \in f_{nkr}} [\lambda_i(t_{mi}) \left\lfloor \frac{\pi_{mi}^w}{\lambda_i(t_{mi})} \right\rfloor], \quad (15)$$

where $\lfloor x \rfloor$ is the operator denoting the largest integer smaller than or equal to x . π_{mi}^w is passenger m 's experienced waiting time at stop i . $\lambda_i(t_{mi})$ is the time headway of transit line at i when passenger m arriving at the stop at time t_{mi} . The equation (15) computes the summation of the failing-to-board waiting times at stop i for all passengers departing his/her origin in departure time interval h .

SOLUTION ALGORITHM

The proposed solution algorithm is based on the CE algorithm for solving general dynamic traffic assignment problem (13, 14, 15). The CE method is a stochastic optimization algorithm for solving combinatorial optimization problems (18, 19). The method associates a stochastic mechanism for generating feasible solutions (samples) and iteratively improves the solution quality based on the performance of the samples. The CE algorithm is a learning algorithm based on the the minimization of the Kullback–Leibler distance (cross entropy) to unknown optimal density (user equilibrium assignment density). We state the basic concept of the CE algorithm and develop the solution algorithm for the DSOR problem.

To illustrate the basic concept of the CE algorithm for traffic assignment problem, we consider the path assignment problem with OD pair k connected by a set of routes R_k . We associate a probability function \mathbf{p}^w with w the iteration index to generate solutions. The performance of the assignment on the paths depends on the experienced travel cost. Let us define the path performance function by Boltzmann distribution (20):

$$H_r(\gamma) = e^{-C_r(\mathbf{t})/\gamma} \quad (16)$$

where γ is a control parameter or temperature. It can be seen that decreasing the value of γ increases the path flows on cheaper paths.

As we seek to obtain a target probability density \mathbf{p}^* , i.e. user equilibrium assignment density, a direct way is generating very large samples by crude Monte-Carlo simulation, which is generally impractical. An alternative way is using importance sampling density \mathbf{p} in

the family of $\{\mathbf{p}(\mathbf{f}; \mathbf{v})\}$, parameterized by some performance information \mathbf{v} . The objective is to get better assignment towards to cheaper paths iteratively based on the path performance information. If a path is found with cheaper cost, its choice probability will be adjusted to have higher flows at the next iteration. To derive the optimal importance sampling density from the current known probability density \mathbf{p}^w , where w is an iteration index, we can solve the cross entropy minimization problem, equivalent to the following maximization problem (18):

$$\max_{\mathbf{p}} E_{\mathbf{p}^w}[H(\gamma) \ln \mathbf{p}] \quad (17)$$

$$\text{st. } \sum_{r \in R_k} p_r = 1, \quad \forall p_r \geq 0, \quad (18)$$

where w is an iteration index.

Note that \mathbf{p}^w is the current known density and \mathbf{p} is the importance sampling density to be derived.

The Lagrangian function of (17)-(18) is written as:

$$L = \sum_{r \in R_k} [p_r^w e^{-C_r/\gamma} \ln \mathbf{p}] + u[\sum_{r \in R_k} p_r - 1] \quad (19)$$

where u is the Lagrangian multiplier.

The first order optimality condition states:

$$\frac{\partial L}{\partial p_r} = \frac{p_r^w e^{-C_r/\gamma}}{p_r} + u = 0 \quad (20)$$

By summing over p_r for $r \in R_k$, we obtain:

$$\sum_{r \in R_k} \frac{p_r^w e^{-C_r/\gamma}}{-u} = 1 \Leftrightarrow \sum_{r \in R_k} p_r^w e^{-C_r/\gamma} = -u = \frac{p_r^w e^{-C_r/\gamma}}{p_r} \quad (21)$$

Hence, we can derive the importance sampling density p_r for iteration $w+1$ as:

$$p_r^{w+1} = p_r^w \frac{e^{-C_r^w/\gamma^w}}{\sum_{s \in R_k} p_s^w e^{-C_s^w/\gamma^w}}, \quad \forall r \in R_k \quad (22)$$

This adjusted probability function (importance sampling density) favours the shifting of flows on shorter paths. The shifting force is determined by the control parameter γ^w , minimized under the constraint that the summation of the changes of probabilities is bounded by a divergent series, i.e.:

$$\text{Min } \gamma^w \text{ subject to } \sum_{r \in R_k} |p_r^{w+1} - p_r^w| \leq \alpha^w \quad (23)$$

where $\alpha^w = \frac{\theta}{w}$ is a numerical divergent series ($\alpha^w \rightarrow 0$ as $w \rightarrow \infty$ and $\sum_{w=1}^{\infty} \alpha^w = \infty$) such

that the flow adjustment converges to fix points. θ is a positive constant.

As the iterative probability update process (22)-(23) makes $\gamma_w \rightarrow \infty$, which makes the field converge to the fixed points. The reader is referred to (13, 14, 15) for more detailed description.

Main Algorithm

For solving the DSOR problem of (3)-(5), the CE method is proposed as follows.

Step 1 (Initialization): generate acyclic multimodal path choice set by the modified k-shortest algorithm (21) or the stochastic route generation approach (22). Initialize uniform probability distributions for path choice.

Step 2 (Dynamic network loading): Loading the time-dependent travel demand on the multimodal transit network and run the transit system simulation. When passengers arrive at his/her respective destination, compute his/her generalized travel cost by (2). Compute the time-dependent path marginal cost with respect to passenger's departure time interval by (14) and (15). Note that given a departure time interval h , departure time is randomly selected within the interval $[t_0 + h\Delta, t_0 + (h+1)\Delta)$ with t_0 is the earliest departure time instant.

Step 3 (Assignment probability update): compute the time-dependent path assignment probability as

$$p_{hkr}^{w+1} = p_{hkr}^w \frac{e^{-\hat{\Gamma}_{hkr}^w / \gamma_{hk}^w}}{\sum_{s \in R_{hk}} p_{hks}^w e^{-\hat{\Gamma}_{hks}^w / \gamma_{hk}^w}}, \quad \forall r \in R_{hk}, \quad (24)$$

where $\hat{\Gamma}_{hkr}^w = \frac{\tilde{\Gamma}_{hkr}^w}{\bar{\Gamma}_{hk}^w}$ is the normalized path marginal cost with respect to h and k . $\tilde{\Gamma}_{hkr}^w$ is estimated

path marginal cost estimated by (14)-(15). The average path marginal cost $\bar{\Gamma}_{hk}^w$ is computed by

$\bar{\Gamma}_{hk}^w = \frac{1}{|R_{hk}|} \sum_{r \in R_{hk}} \tilde{\Gamma}_{hkr}^w$. Note that R_{hk} is the path choice set generated at Step 1. γ_{hk}^w is the control

parameter with respect to h and k resulting from the solution of (23).

Step 4 (Stop criteria): when $w = w^{\max}$ or the resulting probability updates stabilize, stop;

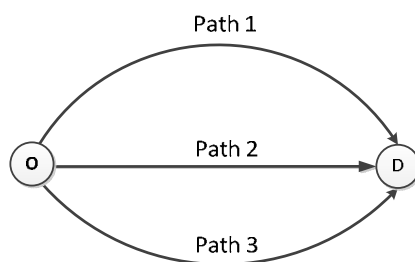
otherwise goto Step 2.

COMPUTATIONAL STUDY

In this section, we present the computational results of the proposed CE method for static system optimal routing and the DSOR problem. First, we test the CE algorithm on static case to validate its computational performance compared with the widely used MSA algorithm and the optimal solution. For dynamic case, we compare the obtained system optimal solutions by both methods under different loading factors. The results are also compared with *user optimal routing* solution to evaluate total system travel cost saving under different scenarios.

Static System Optimal Routing

A simple static network with one OD pair connected by three paths with non-linear path cost functions is depicted in Figure 2. For solving the static system optimal routing problem, we first compute the marginal path cost for every path by (13) and then apply the CE algorithm to solve the VI problem of (9)-(10). The convergence result of the CE algorithm is illustrated in Fig. 3. It shows the CE algorithm finds the system optimal solution with a gap of $8.72e-007$ compared with the optimal solution. Its convergence speed outperforms the MSA method. The total travel cost of the CE algorithm is 229.304 (the optimal solution is 229.3038). The obtained path flow on path 1, 2 and 3 are 2.8390, 4.3117, and 2.8492, respectively. The initial control parameter θ below (23) is set as 0.1.



$$t_1 = 10(1.0 + 0.15(\frac{x_1}{2})^4)$$

$$t_2 = 20(1.0 + 0.15(\frac{x_2}{4})^4)$$

$$t_3 = 15(1.0 + 0.15(\frac{x_3}{3})^4)$$

$$x_1 + x_2 + x_3 = 10$$

FIGURE 2 A small static network example

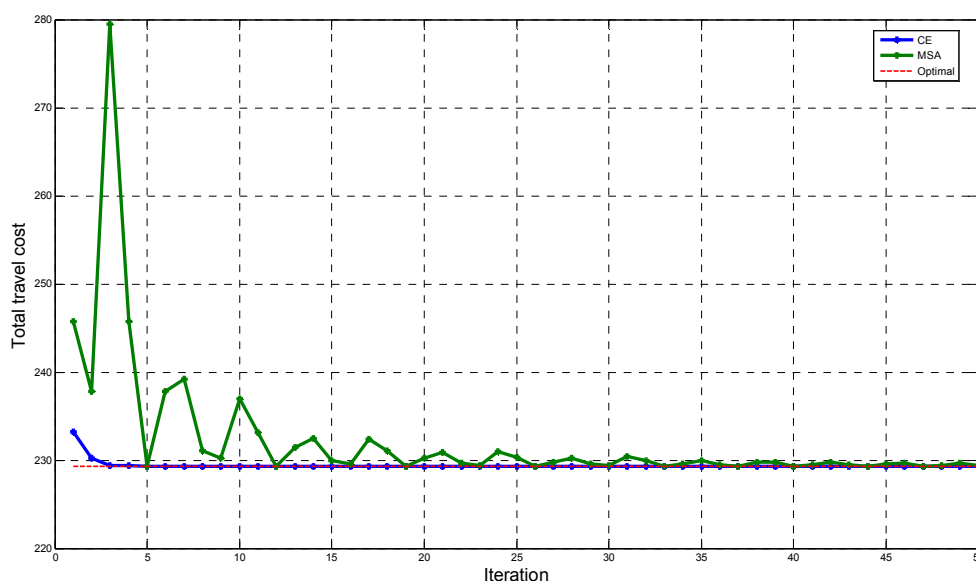


FIGURE 3 Convergence result of the CE and MSA algorithms for static system optimal routing problem

Dynamic System Optimal Routing on Transit System

The experimental study for dynamic DSOR problem is implemented on a small multimodal transit network (Fig. 4). The network is composed of one bus line and two metro lines connecting 4 OD pairs. For simplification, the transit operations are available only in one direction. However, one can extend the transit network and related operations in both directions with little cost. There are totally 19 nodes and 40 links for the multilevel multimodal transit network. For the transit operations setting, the speed of metro and bus are set as 20.0 and 12.5 m/sec, respectively. The capacity of vehicle for metro and bus is set as 200 and 40 passengers/vehicle, respectively. The stop times at stations for metro and bus is set as 20 seconds uniformly. The frequency of metro and bus is set as 20 and 6 vehicles per hour, respectively. We assume that all passengers can board/alight the vehicles within the stop times. The length of boarding, alighting and transfer arcs is set as 100 m for both modes. The access distance between O/D nodes and related metro/bus stations is set as 300 m.

For demand setting, the departure time period is set as 60 minutes with discretized time interval of 5 minutes. There are three origins (node 1, 2, and 3) and one destination (node 4). For each OD pair, the time-dependent demand profile for the reference scenario is set as 160 passengers/20 minutes, 320 passengers/20 minutes and 160 passengers/20 minutes consecutively to generate congestion situation. Four loading factors with respect to the reference scenario are tested in the numerical study, namely 1.0, 1.5, 2 and 2.5 (Table 1). The multiagent transit system is implemented by discrete event simulation technique based on

C++.

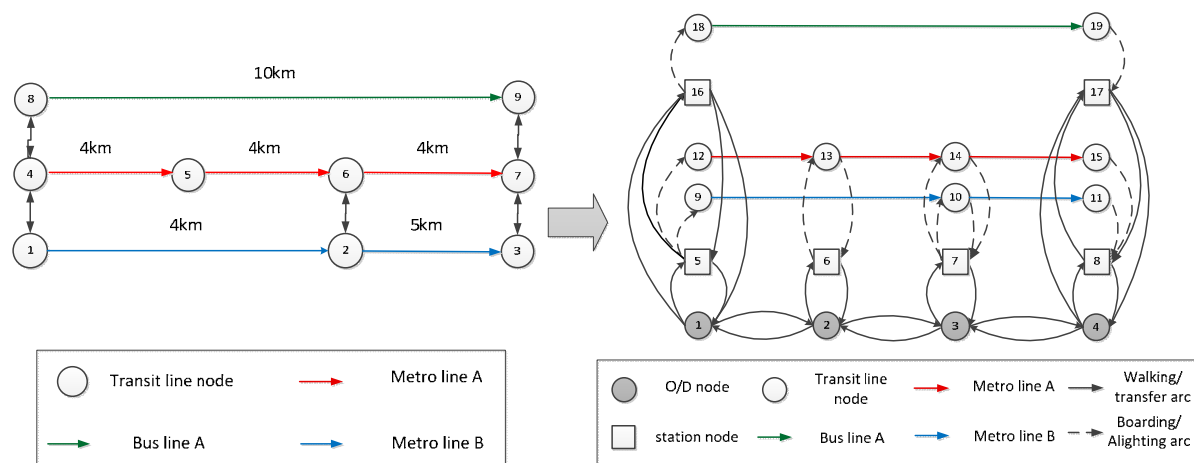


FIGURE 4 A simple transit network with one bus and 2 metro lines (left); Presentation of multilevel multimodal transit network (right)

For the OD pair (1, 4), the path choice set contains 5 paths: a) 1-16-18-19-17-4 (1371 sec.); b) 1-5-12-13-14-15-8-4 (1621 sec.) c) 1-5-9-10-11-8-4 (1381 sec.) d) 1-5-12-13-14-7-10-11-8-4 (1934 sec.); e) 1-5-9-10-7-14-15-8-4 (1534 sec.). For OD pair (2, 4), there is only one path: 2-6-13-14-15-8-4 (1301 sec.). For the OD pair (3, 4), there is two paths: a) 3-7-14-15-8-4 (981 sec.); b) 3-7-10-11-8-4 (1061 sec.). Note that the path travel time in the parentheses includes the average waiting time for boarding the transit vehicle.

The typical convergence pattern of the CE and MSA method for the DSOP problem is shown in the Figure 5. The result indicates that the two algorithms converge to near-optimal solution after 5 iterations. The MSA method has a higher total travel time at initial iteration due to the assignment of passengers' flows on the initial shortest path.

The computational results of the CE algorithm for the DSOR problem are shown in Table 1. The result indicates that the CE method performs better than the MSA method for solving the DSOR problem and dynamic user optimal (UE) routing problem for most cases. As expected, when the loading factor increases, the total travel time differences between the system-optimal and user-optimal routing increases accordingly. This is due the fact that as the level of congestion increases, some user may be assigned to longer travel route to reduce total system travel time. The result is on the line with previous study (10). Moreover, when the loading factor increases, the percentage of total time savings between the DSOR and UE solutions becomes higher from 0.20% (loading factor equals 1) to 0.33% (loading factor equals 2.5). As the possible routes for re-assigning passengers are relatively limited, the percentage increasing of total time savings becomes not very significant when the system is highly congested. However, it can be expected that when more routes can be used to each destination, the percentage of total time savings might become more significant between the DSOR and UE routing strategies. Moreover, the limited difference between the DSOR and UE

routing strategies may be due to underestimating the time-dependent path marginal cost. If the scheduled delay cost of early/late arrival at destination is considered in the generalized cost computation, more significant difference could be expected.

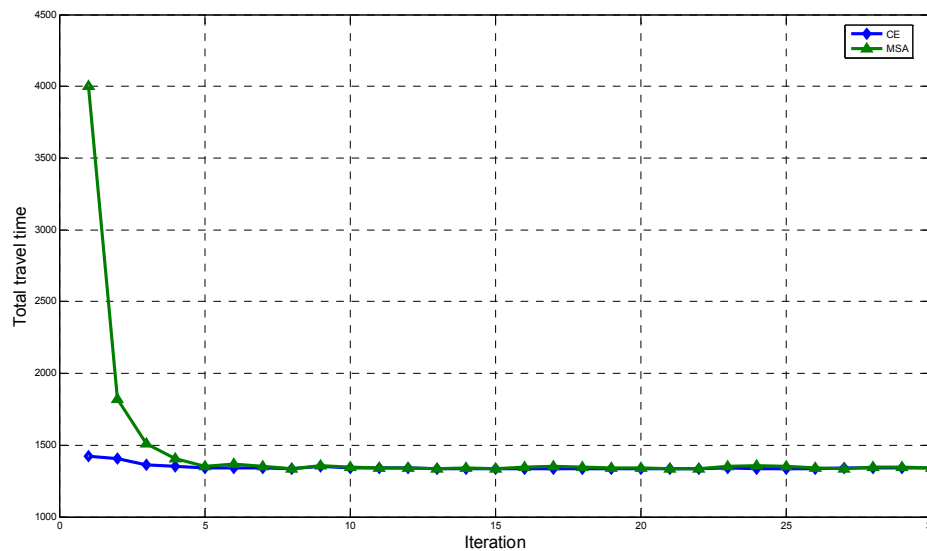


FIGURE 5 The typical convergence pattern of the CE and MSA method for the DSOR problem

TABLE 1 The computational results of the CE and MSA methods for the DSOR problem

Loading factors	Number of passengers	Dynamic system optimal routing (SO)				Total travel time difference between SO and UE solution (hr)
		CE	MSA	Best	Average travel time per passenger (sec.)	
		Total travel time (hr)	Total travel time (hr)	Total travel time (hr)	Average travel time per passenger (sec.)	
1	1920	659.06	659.10	659.06	1235.7	1.29 (0.20%)
1.5	2880	993.45	994.30	993.45	1241.8	3.08 (0.31%)
2	3840	1334.22	1329.39	1329.39	1246.3	3.24 (0.24%)
2.5	4800	1672.70	1671.96	1671.96	1254.0	5.53 (0.33%)
		User optimal routing (UE)				
		CE	MSA	Best	Average travel time per passenger (sec.)	
		Total travel time (hr)	Total travel time(hr)	Total travel time(hr)	Average travel time per passenger (sec.)	
1	1920	660.36	661.29	660.36	1238.2	
1.5	2880	996.53	998.55	996.53	1245.7	
2	3840	1332.63	1333.33	1332.63	1249.3	
2.5	4800	1677.49	1681.21	1677.49	1258.1	

Remarks: The result is based on the solutions obtained after 20 iterations. The parameter θ below (25) is set within [1.70, 1.8] for all tested scenarios.

As for the influence of θ (below Eq. (23)) on the performance of the CE method for the DSOR problem, the result is shown in Figure 6. It illustrates that the initial value of the parameter influences the obtained solution quality. To find better solution quality, one can use line search techniques to obtain optimal θ to the problem to be solved.

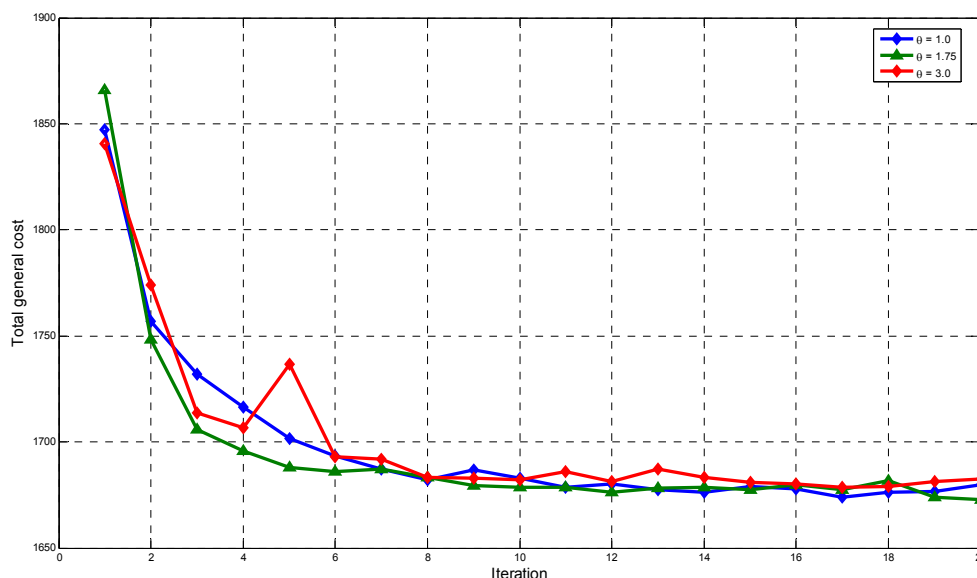


FIGURE 6 Influence of θ on the solution quality obtained by the CE method for the DSOR problem

CONCLUSIONS

The design of multimodal itinerary guidance systems aims at providing passengers with efficient route guidance to arrive to destination. From the system point of view, it is desirable to reduce total passengers' travel times. However, traditional time-dependent shortest-path-based route guidance system might be inefficient due to the service capacity constraints. In this paper, we propose a dynamic system optimal routing model on a multimodal transit network. The network is modeled by a multilevel directed graph to simulate explicitly passengers' movements and transit vehicle operations. The dynamic system optimal problem is formulated as a variational inequality problem for dynamic user equilibrium in terms of time-dependent path marginal cost. The computation of time-dependent path marginal cost is based on the passengers' additional waiting times due to failing to board the arriving vehicles. We propose the cross entropy method for solving the dynamic system optimal problem. The computational results on static cases and dynamic cases show that the proposed algorithm performs better than the method of successive averages. We compare also the system optimal and user optimal routing strategies with respect to different network loading factors. The numerical result on a small network suggests that when the congestion level of transit system increases, the system optimal routing may reduce

total system travel time.

Further study includes some numerical experiments on realistic multimodal transit networks and the computation of the user-constrained multimodal shortest paths. Moreover, the incorporation of more realistic transit system operation modeling and passenger route choice behaviour modeling for non-guided users are desired. It is also interesting to test the scenarios with different proportion of informed users of the route advisory system.

REFERENCES

1. Foo, H. M., H. M. Lenong, Y. Lao, and H. C. Lau. A multi-criteria, multi-modal passenger route advisory system. In: Proceedings of the IES-CTR, Singapore 1999.
2. Lo, H. K., C. W. Yip, and B. Mak, Passenger route guidance system for multi-modal transit networks. *Journal of advanced Transportation*, Vol. 39, No. 3, 2005, pp.271–288.
3. Galvez-Fernandez, C. D. Khadraoui, H. Ayed, Z. Habbas, and E. Alba. Distributed Approach for Solving Time-Dependent Problems in Multimodal Transport Networks. *Advances in Operations Research*, Vol. 2009, Article ID 512613, 15 pages, 2009. doi:10.1155/2009/512613.
4. Weymann, J., J. L. Farges, and J. J. Henry. Optimization of traffic dynamic route guidance with drivers' reactions in a queue-based model. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 25, No 7, 1995, pp. 1161-1165.
5. Jahn, O., R. Möhring, A. Schulz, and N. S. Moses. System-Optimal Routing of Traffic Flows with User Constraints in Networks with Congestion. *Operations research*. Vol. 53, No. 4. 2005, pp. 600–616.
6. Bottom, J. A. 2000. *Consistent anticipatory route guidance*. Ph.D. thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.
7. Chow, A. H. F. Properties of system optimal traffic assignment with departure time choice and its solution method. *Transportation Research Part B*, Vol. 43, 2009, pp. 325–344.
8. Zuurbier, F. S. *Intelligent Route Guidance*, PhD Dissertation, 2010, Delft University of Technology
9. Beckmann M., C. B. McGuire, and C. B. Winsten. Studies in the economics of transportation. 1955: Yale University Press, New Haven, Connecticut; also published as Rand-RM-1488-PR, Rand Corporation, Santa Monica.
10. Peeta, S., and H. Mahmassani. System optimal and user equilibrium time-dependent traffic assignment in congested networks. *Annals of Operations Research*, Vol. 60, No. 1, 1995, pp. 80-113.
11. Powell, W. B., and Y. Sheffi. The Convergence of Equilibrium Algorithms with Predetermined Step Sizes. *Transportation Science*, Vol. 16, 1982, pp. 45-55.

12. Tong, C. O., and S. C. Wong. A predictive dynamic traffic assignment model in congested capacity-constrained road networks. *Transportation Research Part B*, Vol. 34, No. 8, 2000, pp. 625–644.
13. Ma, T.-Y., and J.P. Lebacque. A cross entropy based multi-agent approach to traffic assignment problems. In: Appert-Rolland, C., Chevoir, F., Gondret, P. Lassarre, S., Lebacque, J.P., Schreckenberg, M. (Eds.), *Proceedings of the Traffic and Granular Flow '07*, Springer Berlin Heidelberg, 2007, pp. 161-170.
14. Lebacque J.-P., T.-Y. Ma, and M. M. Khoshyaran. The cross-entropy field for multimodal dynamic assignment. In: *Proceedings of the Traffic and Granular Flow '09*, 2009, Springer (to appear)
15. Ma, T.-Y., and J.P. Lebacque. A cross-entropy based multiagent approach for multimodal activity chain modeling and simulation, *Transportation research Part C*, 2012, <http://dx.doi.org/10.1016/j.trc.2011.11.014>.
16. Ma, T.-Y. A hybrid multiagent learning algorithm for solving the dynamic simulation-based continuous transit network design problem. In: *Proceedings of 2011 Conference on Technologies and Applications of Artificial Intelligence*, IEEE Computer Society Conference Publishing Services, 2011, pp. 113-118.
17. Ma, T.-Y. and J. P. Lebacque. A dynamic packet-based multi-agent approach for large scale multimodal network simulation. CD-ROM, In: *Proceeding of 10th International Conference on Application of Advanced Technologies in Transportation*, Greece, 2008.
18. Rubinstein, R. Y. The cross-entropy method for combinatorial and continuous optimization. *Methodology and computing in applied probability*, Vol. 1, No. 2, 1999, pp. 127-190.
19. De Boer, P.T., D. P. Kroese, S. Mannor, and R.Y. Rubinstein. A Tutorial on the Cross-Entropy Method. *Annals of Operations Research*, 134, 2005, pp. 19–67.
20. Helvik, B. E., and O. Wittner. Using the Cross-Entropy Method to Guide/Govern Mobile Agent's Path Finding. *Networks. Lecture Notes in Computer Science*, 2001, Vol. 2164, 2001, pp. 255-268.
21. Ma, T.-Y. A path-based cross entropy approach for solving dynamic multimodal transit assignment problem. Paper presented in the 13th International Conference of the International Association for Travel Behaviour Research (IATBR), Toronto, Canada, 2012
22. Bliemer, M., and H. Taale. Route generation and dynamic traffic assignment for large networks. In: *Proceedings of the first International Symposium on Dynamic Traffic Assignment*, 2006, Leeds, UK.