



HAL
open science

a cross-entropy based multiagent approach for multiclass activity chain modeling and simulation

Tai-Yu Ma, Jean-Patrick Lebacque

► To cite this version:

Tai-Yu Ma, Jean-Patrick Lebacque. a cross-entropy based multiagent approach for multiclass activity chain modeling and simulation. 2011. halshs-00310903v3

HAL Id: halshs-00310903

<https://shs.hal.science/halshs-00310903v3>

Preprint submitted on 11 May 2011

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.

A cross entropy based multiagent approach for multiclass activity chain modeling and simulation

Tai-Yu Ma^a, Jean-Patrick Lebacque^b

^a Laboratoire d'Economie des Transport (LET), LET-ISH, 14, Avenue Berthelot F-69363 Lyon Cedex 07

^b Institut français des sciences et technologies des transports, de l'aménagement et des réseaux (IFSTTAR), Le Descartes 2, 2 rue de la Butte Verte 93166 Noisy-le-Grand Cedex

Abstract

This paper attempts to model complex destination-chain, departure time and route choices based on activity plan implementation and proposes an arc-based cross entropy method for solving approximately the dynamic user equilibrium in multiagent-based multiclass network context. A multiagent-based dynamic activity chain model is developed, combining travelers' day-to-day learning process in the presence of both traffic flow and activity supply dynamics. The learning process towards user equilibrium in multiagent systems is based on the framework of Bellman's principle of optimality, and iteratively solved by the cross entropy method. A numerical example is implemented to illustrate the performance of the proposed method on a multiclass queuing network.

Keywords: dynamic traffic assignment, cross entropy method, activity chain, multiagent, Bellman equation

^a Corresponding author. Tel. : +33 4 72 72 64 46 ; Fax : +33 4 72 72 64 48
Email addresses : tai-yu.ma@let.ish-lyon.cnrs.fr (T.Y. Ma), lebacque@inrets.fr (J.P. Lebacque)

1. Introduction

The dependency of travel demand with activity planning and location choice is well recognized. Travelers usually organize their trip-chain such that the required travel time or cost is minimized for connecting their activity destinations. As travel demand is derived from activity realization, a variety of activity-based approaches have been proposed in the literature (see Jones et al., 1990). These studies tried to model travel demand by considering related decision choices with respect to activity realization. However, most of them focused either on static analysis of trip-chain characteristics or on single activity-based travel demand analysis. The complete dynamic activity chain modeling and simulation has been less studied and still is a difficult and complicated research issue in transportation science.

The analysis of activity chaining behavior, originated from geography and urban planning in the early 1970s, focuses on the analysis of activity chain characteristics and its relationship with socio-demographic factors. These studies provided the analysis framework of activity chain formulation and incited more researchers for activity chain modeling. In a large variety of activity-chain-based travel demand analyses, most of them are based on static analytical framework to derive individual's travel/activity decisions based on time-space constraints, neglecting traffic flow dynamics. These studies are mainly based on mathematical programming approaches (Recker, 2001), computational process modeling approaches (Arentze and Timmermans, 2004) or static user equilibrium framework (Maruyama and Sumalee, 2007).

For dynamic activity chain modeling and simulation, Axhausen (1990) reported an activity chain simulation model combining activity-chain-based travel decision choice model and mesoscopic traffic flow propagation model. Lam and Yin (2001) proposed a dynamic activity-based traffic assignment model combining activity and route choice. The equilibrium is formulated as a variational inequality problem for which a network loading procedure combining the method of successive averages and Frank-Wolfe algorithm is proposed to obtain approximate solutions of the dynamic user equilibrium. However, the traffic dynamics is simplified by applying the BPR (Bureau of Public Road) type function, and the activity chaining problem is not treated. Ramadurai and Ukkusuri (2008) proposed an integrated activity-based model for destination, starting time of activity, duration of activity and route choice. A first order macroscopic traffic flow model is utilized to capture traffic flow dynamics. The decision choice is modeled on the basis of utility maximization by summing the utility of activities and the disutility of trips conducted in a travel-activity sequence. A day-to-day route flow adjustment process based on Euler's method is proposed to obtain an approximate of dynamic user

equilibrium (DUE). However, the proposed route-based adjustment process is problematic when the size of network becomes large.

Recently, increasing applications of the multiagent framework for dynamic activity chain modeling and simulation have been proposed (Cetin et al., 2002; Raney et al., 2003; Rieser et al., 2007). The multiagent framework is very convenient to simulate complex transportation systems composed of a large number of homogeneous / heterogeneous individuals (agents) interacting between them and with their environment. In these models, each traveler is modeled as an autonomous agent aiming to implement a sequence of planned activities on a daily basis. The departure time and route choice are undertaken in a sequential and heuristic way, implying that travelers search satisfaction solutions under dynamical environment. Traveler's learning process is modeled in an iterative day-to-day adjustment process based on the experienced performance of travel and activity choices. However, there are still few studies on efficient algorithms for computing approximate solutions of DUE on such a non-cooperative multiagent system. For this issue, existing convergence results and conditions for convergence towards the Nash equilibrium in multiagent systems have been proposed in simple static or general sum stochastic games (Fudenberg and Levine, 1997; Hu and Wellman 2003; Busoniu et al., 2008). However, the general conditions allowing multiagent systems converge towards a Nash equilibrium require that each agent has complete information of the strategic profiles, i.e. payoffs of action/choice of the other agents. This condition is generally unavailable in multiagent-based dynamic traffic assignment case. Different with multiagent learning algorithms developed in artificial intelligence, Ma and Lebacque (2007) proposed a cross entropy method solving the static/dynamic traffic assignment problem in multiagent transportation systems. It has been proved that the proposed cross entropy method converges to fixed points of the cross-entropy field (Ma 2007; Lebacque et al., 2009). The numerical study has shown that the proposed method can find approximations of Nash equilibriums in dynamic environment.

The development of algorithms for solving DUE has been an active research issue in transportation science (see the review of Peeta and Ziliaskopoulos, 2001). The formulation of DUE can be based on variational inequalities, fixed-point or non-linear complementarity problem (Patriksson, 1994). Several solution techniques have been proposed for solving DUE: projected dynamical system approaches (Bertsekas and Gafni, 1982; Nagurney 1993), dynamical system approaches (Smith, 1979, 1993), and the method of successive averages (Tong and Wong, 2000). Dafermos (1972, 1980, 1982) proposed a series of papers in studying traffic equilibrium problem in multiclass transportation network with asymmetric link cost functions. The solution algorithms and the necessary conditions for the convergence of the algorithm have been proposed under the monotonicity property of travel cost. However, for multiclass / multimodal

simulation-based DUE problem, the difficulty in obtaining approximations of DUE remains on the non-monotonicity property of travel cost when more than one class of users are incorporated in a link (Wynter, 2001).

This paper suggests an activity-chain-based dynamic traffic assignment model on the basis of traveler's activity program realization. The travel choice concerns destinations, departure time and route choice. An extended point queue model is proposed to capture traffic flow dynamics in network. For demand modeling, a multiagent approach is developed for traveler's complex activity chain modeling and simulation. Different with traditional trip-based approaches, the proposed activity-based approach considers that travel demand is derived from the realization of an activity program (plan). The complex activity-chain-based DUE condition is formulated by applying dynamic programming approach. Each traveler is considered as a bounded-rational user (utility-maximiser) aiming, at each activity location, at maximizing the total expected gain of net activity value (utility) for not yet implemented activities in his/her activity program. As the environment is stochastic in the presence of multiagent interactions, traveler's learning process is indeed similar to multiagent reinforcement learning process in distributed artificial intelligence. The solution of optimal travel choice at each decision-making stage is based on the idea of Bellman's principle of optimality, resulting in an approximate of dynamic user equilibrium with respect to total net activity values obtained from activity program realization. According to the experience of all travelers, the system learns iteratively to shift travelers to more attractive travel choices. As link costs are asymmetric with the presence of multiclass users, the traditional derivative-based method is not applicable. Hence, an arc-based cross entropy (CE) method (Rubinstein, 1999; Helvik and Wittner, 2001; Ma and Lebacque, 2007) is proposed to approximate activity-chain-based DUE. The basic idea is that we utilize a set of probability distributions guiding travelers towards DUE, which is considered as a rare event among all possible assignments. The DUE as a rare event is achieved by iteratively updating the probability distributions based on minimization of cross entropy. The CE method is a derivative-free method convenient for solving asymmetric multiclass dynamic traffic assignment problems. It has also been shown to be more efficient than the dynamical system approach since it optimizes the adjustment step size towards unique/multiple equilibrium points. (Ma and Lebacque, 2007; see also Jin, 2007; Smith 1979, 1993 for alternative dynamical systems approaches). Different with the algorithms based on route flow adjustment procedure, the proposed approach is arc-based, avoiding the difficulty of route enumeration for all OD pairs.

The remainder of this paper is organized as follows. First, the basic assumptions of the proposed dynamic activity chain model are presented. Then activity value measurement and activity program settings are discussed. We propose an extended point

queue model to capture traffic flow dynamics on road network and discuss the modeling approach for multiclass network. The approximate solution of DUE condition on the basis of activity chain realization is obtained based on the idea of Bellman's principle of optimality, solved iteratively by cross entropy method. Having formulated the DUE condition on individual basis, an arc-based CE approach is proposed to approximate simulation-based multiclass DUE. A numerical example is illustrated on a bimodal road network to illustrate the performance of the solution algorithm. Finally, the conclusions and future extensions are discussed.

2. Model formulation

The basic assumptions of the proposed activity chaining model are discussed as follows:

(i) we consider a multiclass road network with a set of behaviorally homogeneous travelers, i.e. traveler's behavior depends only on the net value of activities (accounting for only travel cost and economic value of activity). This simplification allows us to investigate the performance of the solution algorithm without loss of generality. The current model can be easily extended by classifying users into different behaviorally homogeneous groups. Each traveler is modeled as an *agent* aiming to implement his/her activity program (plan) for a period of time. The activity program is simplified as a sequence of activities characterized by desired starting times assumed known *a priori* and fixed. The locations of two consecutive planned activities are assumed to be different and need a trip linkage. We limit our analysis only to the destination-chain, departure time and route choice problem. The mode choice in trip chain is not taken into account in this study. (ii) The activity model is based on *Accessibility to Vacant Activities* (AVA) model (Leurent, 1999; Ma and Lebacque, 2006). Activity demand is located at origins, and activity supply at destinations. The activities are assumed countable with constrained capacity (fixed activity supply) for which they possess different gross economic values, identically perceived by all individuals. Also, we suppose each individual chooses one best vacant activity at destination for which occupied activities cannot be served by the other individuals. As the activity value is generally difficult to measure, we assume that its value follows some probability distributions with differentiable density. However, more elaborated activity value functions taking into account activity duration can also be applied (Lam and Yin, 2001). (iii) For the implementation of activity program, it is assumed that each traveler aims to perform the totality of his scheduled activities and maximize his/her total net activity value. Route cost is assumed arc-additive. However, route-specific cost can also be added. Travel decisions are undertaken with respect to destination chain, departure time and route choice. The two latter decisions are considered in a sequential manner for each scheduled activity. Travelers are assumed to have no priori information about traffic

condition and activity value distribution at destinations. (iv) The departure time choice is considered within a reasonable time period, i.e. between actual arrival time of the destination of current activity and the desired starting time of next planned activity. The reasonable time period is discretized into a set of equal time interval. Travelers choose a departure time interval and then a random departure time instant is selected. Travelers aim to arrive to activity destinations on expected arrival time, and an early/late arrival penalty is imposed if the actual arrival time differs from the expected arrival time.

The traveler's activity chaining behavior is summarized as follows. Let each traveler have a fixed activity program to engage in a period of time. Keeping planned activities in mind, each traveler makes his/her destination-chain choice at the origin and then conducts sequentially departure time and route choice for the next activities in his/her activity program. The optimal choice decision at each decision stage is the solution which maximizes the expected net activity values for not yet implemented activities. The system learns the optimal strategy iteratively based on the average performance of travel alternatives, measured by net activity values, i.e., gross values of activity minus its generalized travel cost. It should be noted that traveler's activity program doesn't change from one day to another. The system searches a long term user equilibrium state in an activity chaining context. Although the proposed model simplifies the traveler's complex decision making, it captures the essential of travel-cost/activity-value tradeoff in the activity program realization process.

2.1 Activity program and economic value of activity chain

Let an activity program g be defined as a sequence of planned activities that a traveler would like to engage during a period of time, i.e.:

$$g = \{0, 1, 2, \dots, i, \dots, n\} \quad (1)$$

where i denotes the i th planned activity. Each activity is characterized by its desired starting time τ and its economic value. The sequence of desired starting times of activities satisfy the condition $\tau_0 < \dots < \tau_n$. Let O be the set of origins, and D the set of destinations of activities. The set of possible OD pairs is denoted by $K = \{(i, j) | \forall (i, j) \in \{O \cup D\}\}$. A route-chain of activity program g with origin o is defined as a sequence of routes linking all scheduled activity destinations:

$$u = \{r_1, r_2, \dots, r_n\}, \quad \forall u \in U_g^o, \quad \forall r_i \in R_{k_{i-1,i}} \quad (2)$$

where o is origin, and r_i is the route connecting two consecutive activities $i-1$ and i . U_g^o denotes the set of route-chains for activity program g with departure origin o , and $R_{k_{i-1,i}}$ the set of routes connecting OD pair k of activity $i-1$ and i .

We now describe briefly the AVA activity model. Let A_d represent the number of vacant activities at destination d , and $T_d(t)$ the number of served activities at destination d at time t , $0 \leq T_d(t) \leq A_d$. Its economic value with respect to the destination d follows an exponential probability distribution function as:

$$h_d(v) = \lambda_d \exp(-\lambda_d(v - m_d)), \forall d \in D \quad (3)$$

where $v \geq m_d$ and $\lambda_d \geq 0$.

Let $H_d(v) \equiv \int_{-\infty}^v h_d(v)$ denote the cumulative function of activity value v , and $H_d^{-1}(v)$ its inverse function. The activity value $v(t)$ obtained by individuals when arriving at destination d at time t is calculated as $v(t) = H_d^{-1}(1 - \frac{T_d(t)}{A_d})$.

By assuming the additivity of economic value, the net activity value obtained at destination is calculated as its gross activity value minus its generalized route cost. The later is composed of two parts, one being its trip time and the other the early/late arrival penalty with respect to desired starting time of activity. Based on bounded-rational behavior assumption, an indifference interval is introduced to reflect the tolerable schedule variation. We use a piecewise early/late arrival penalty function to reflect the arrival cost associated to the ideal arrival time (Vickrey, 1969; Mahmassani and Chang, 1985). The generalized trip cost of the route r_i , connecting the destinations of scheduled activity $i-1$ and i , is formulated as

$$C_{r_i}(t) = \pi_{r_i}(t) \times \mu + \mu_\alpha \times \max(0, \tau_i - \Delta - t_{d_i}^{\text{arr}}) + \mu_\beta \times \max(0, t_{d_i}^{\text{arr}} - \tau_i - \Delta) \quad (4)$$

where $\pi_{r_i}(t)$ is the travel time of route r_i when entering the initial point of the route at time t , $t_{d_i}^{\text{arr}}$ the arrival time at the destination d_i of activity i . μ is the unitary economic value of travel time. μ_α and μ_β are unitary penalty associated with early and late arrival, respectively. Δ is the half of tolerable schedule delay interval without penalty. Based on the experimental result (Small, 1982), we assume that the condition $0 < \mu_\alpha < \mu < \mu_\beta$ holds. Note that the penalty value is evaluated by traveler's actual arriving time to destination, depending on his/her departure time and traffic condition. The penalty value reflects traveler's experienced performance of departure time and route choice, according to which travelers adjust their departure time and route choice on next iteration. This learning process is reflected in the related choice probability update calculation ((19)-(21)).

The net economic value of activity i obtained by a traveler at destination depends on the traveler's: (1) leaving time t_{i-1} from previous activity $i-1$; (2) arrival time t_i to the next activity i ; (3) route choice r_i . The net activity value is calculated as:

$$v_i^*(r_i, t_i, t_{i-1}) = v_i(t_i) - C_{r_i}(t_{i-1}) \quad (5)$$

where $v_i(t_i)$ is the gross economic value of activity i obtained at destination when arriving at time t_i . $C_{r_i}(t_{i-1})$ is the generalized travel cost of route r_i when entering the initial point of the route at time t_{i-1} .

Based on (5), the net activity value of activity chain $v_{og}(u, t)$ is the sum of net activity values received at destinations with respect to activity program g , i.e.

$$v_{og}(u, t) = \sum_{\forall i \in g} v_i^*(r_i, t_i, t_{i-1}), \quad \forall g, \quad \forall o \quad (6)$$

where u denotes a trip chain. t is the entering time of transport network.

2.2 Dynamic traffic flow propagation

The network is described as a graph $G = (V, E)$ with a set of nodes V and a set of directed links E . The departure times of travelers from their respective origins are chosen by the travelers. The total OD demand D_o is given, and the conservation of the flow at origins applies:

$$D_o = \int_0^T d_o(t) dt, \quad \forall o \in O \quad (7)$$

where D_o is total demand from origin o and $d_o(t)$ the demand flow rate from origin o at time t .

At each destination, an available activity capacity constraint N_d applies:

$$\sum_{e \in M(d)} \int_0^{\bar{T}} x_e(t) dt \leq N_d, \quad \forall d \in D \quad (8)$$

where $x_e(t)$ denotes the flow rate over link e at time t , \bar{T} the time of last vehicle leaves the network. $M(d)$ is the set of entering links of destination node d .

For simplification, each road-user traveler is represented as a road-vehicle agent moving in the road network. Agents chose their routes by choosing links at each node, following a stochastic choice model as described in Section 3. The link traffic flow

model on road network is based on the point queue concept (Kuwahara and Akamatsu 1997). For each link e , the link travel time is calculated as a function of the difference between the arrival and departure curves at arrival time t :

$$T_e(t) = D_e^{-1}(A_e(t)) - t \quad (9)$$

where $D_e(t)$ denotes the cumulative number of vehicles entering link e at time t and $A_e(t)$ the cumulative number of vehicles leaving link e at time t .

We assume that traffic propagation behaviors are different according the traffic conditions (free flow/congestion). The basic model has been adapted in order to take into account intersections in a more realistic way. For links entering an intersection, groups of lanes are distinguished. Each group of lanes corresponds to an outgoing link of the intersection. For links with only one lane or one outgoing link, the regular point queue model is applied. As agents chose an outgoing link, they are added to the corresponding group of lanes. When time-dependent traffic demand exceeds its time-dependent capacity, a traffic queue is generated. Each group of lanes admits its own queue, with its own dynamics in terms of queue generation and dissipation, and First-In-First-Out (FIFO) discipline is respected within each group of lanes. Queues in one group of lanes are assumed not to influence the fluidity of another group of lanes within the same link.

This model can easily be improved at little cost, by implementing a particle discretization of macroscopic traffic flow models (Mammar 2006; Khoshyaran and Lebacque 2008), such as first order models (Daganzo, 1994 ; Elloumi et al., 1994; Buisson et al., 1995 ; Lo, 1999) or second order models (Garavello and Piccoli, 2006; Lebacque et al., 2007).

We consider point-wise intersections, with no specific passing time nor any storage space. However, the intersections can be modeled more realistically by endowing them with physical characteristics, either node supply and demand functions, or internal state dynamics (Lebacque and Khoshyaran 2005, Khoshyaran and Lebacque 2009).

Now let us describe how traffic passes intersections with various entering (upstream) links and various outgoing (downstream) links. The flows from upstream links to downstream links through the intersection are bound by dynamic inflow and outflow capacity constraints, which result from traffic supply and demand constraints. In the congested case, vehicles' link delay depends on the number of vehicles in front of the agent in the same queue. When vehicles from several ingoing links compete for a given outgoing link, the entering order to the chosen outgoing link is given with a probability of choice. This probability expresses how conflicts between agents are resolved in the node. Different rules, based on node optimization models or supply split equilibrium models, could be used (Lebacque and Khoshyaran 2005).

Let us describe briefly the traffic flow constraints for the diverge and merge case in intersections with multiple upstream links e_r and multiple downstream links e_s .

Diverge constraints

In the diverge case, the time-dependent departure flow capacity of lane group l_s within upstream link e_r for the next chosen downstream link e_s is constrained by (supply/demand approach (Lebacque 1996; Lebacque and Khoshyaran 2005):

$$\delta_{e_r}^{l_s}(t) = \min(\alpha_{rs}^{l_s} \delta_{e_r}^*, \sigma_{e_s}(t)) \quad (10)$$

where

l_s is the lane group within upstream link e_r for the next chosen downstream link e_s

$\delta_{e_r}^{l_s}(t)$ is the link outflow from upstream link e_r to downstream link e_s at time t

$\alpha_{rs}^{l_s}$ is a split coefficient corresponding partial outflow capacity from upstream link e_r to downstream link e_s by lane group l_s .

$\sigma_{e_s}(t)$ is the supply of the downstream link e_s at time t

$\delta_{e_r}^*$ is the capacity of the upstream link e_r .

Merge constraints

For intersection i , the total flow entering the downstream link e_s cannot exceed its total time-dependent inflow capacity:

$$\sum_{e_r} \delta_{e_r}^{l_s}(t) \leq \sigma_{e_s}(t), e_r \in M(i) \quad (11)$$

with the same notations as in (10). $M(i)$ is the set of entering links of node i .

These constraints can be managed by random entrance or point wise node models, as mentioned above.

The present multiclass road network setting can be extended to a multimodal transportation system. Buses or other public transport vehicles can be introduced in the model as special agents who can act as moving obstacles in the traffic flow. Buses are also stocked in the point queues at interactions. Metro / tram / train lines could be added in a similar way with capacity constraints and specific operation settings. The connections to road network can be assured via transfer networks. The congestion in metro / tram / train can be modeled as delays on lines and waiting time at stations. The reader is referred to Ma and Lebacque (2008), and Meschini et al. (2007) for more detailed description. The congestion model (the same as the one used for the traffic model) can be viewed as a Lagrangian discretization of a point queue model with a

supply-demand node model (Khoshyaran and Lebacque 2008). The traffic model will be consistent. Travel time estimations are used for activity chain choice (travel choices in activity chain realization) only and evaluated from the agents' trips.

In the multimodal setting, there are multiple equilibriums. The multiagent approach developed in this paper thus aims to construct a plausible equilibrium by a method which is essentially stochastic and which emulates a day-to-day learning process.

2.3 Dynamical user equilibrium based on activity chain

We consider a long-term predictive dynamic user equilibrium given travelers' fixed activity programs. The activity-chain-based dynamical user equilibrium can be stated as an extension of the Wardrop principle (Wardrop, 1952). Hence, for travelers of the same class, i.e. the same origin, mode availability and activity program, the expected total net activity value at each decision making stage, resulting from destination, departure time and route choice, is equal and no less than that of unused choice alternatives. Different with utility maximization, the proposed activity-chain-based model implies that traveler's optimal decision rule is based on the bounded rationality assumption. Travelers aim to maximize future expected net activity value gains under traffic propagation and activity availability dynamics. The dynamic user equilibrium problem in question is essentially the same as a sequential decision making problem for which each traveler tends to find a non-cooperative optimal decision for limited resources. This problem has been recognized as difficult to solve and been remaining an active research issue in the domain of artificial intelligence, computer science and economics (Littman, 1996). We believe there exists such dynamic user equilibrium for the proposed activity-chain model because it is the extension of origin-destination-based dynamic user equilibrium problem in multiagent context (Ma and Lebacque 2007; Lebacque et al., 2009). Further, the problem of determining an equilibrium within the framework of the simulation-based DUE can be reduced by discretization to determining a fixed point of a continuous map in a compact finite-dimensional space.

Different with classical variational inequality formulation of trip-based DUE, an equivalent mathematical statement, on the basis of traveler's decision choice, is based on the idea of Bellman's principle of optimality. As mentioned earlier, the problem to be solved for each agent is the maximization of the expectation of net activity value of activity program under incomplete information. Let the state variable $s(t)$ describe traveler's state at time t , representing the current node of route and the current activity in the activity chain. The decision variable is destination-chain and/or departure time and route choice at each decision stage. Let t_{i-1} be the time of the traveler's arrival to the

destination of activity $i-1$. For travelers at any activity $i-1$, $\forall i \in g$ the optimal travel choice $a_i^*(t_{i-1})$ for next activity i is based on the solution of maximizing the expected net activity values summing from current state to the end of activity program completion. Thereby, for any $i \in g$ it implies:

$$a_i^*(t_{i-1}) = \arg \max_{a_i \in \Gamma(t_{i-1})} E[\sum_{\varphi \geq i} v_\varphi^*(r_\varphi, t_\varphi, t_{\varphi-1})] \quad (12)$$

where

$a_i^*(t_{i-1}) = (r_i^*, t_i^{\text{dep}})$ is the optimal travel choice for next activity i at time t_{i-1} with respect to route choice r_i^* and departure time t_i^{dep} for next activity i .

$\Gamma(t_{i-1})$ is the feasible travel choice set at time t_{i-1} with respect to departure time and route choice.

$v_\varphi^*(r_\varphi, t_\varphi, t_{\varphi-1})$ denotes the net activity value of activity φ , evaluated by (5), depending on selecting route and activity supply dynamics when arriving at destination of φ .

The choice of alternative in the search for the next activity can be viewed as probabilistic. The Bellman's principle of optimality proposes a local solution for the maximization problem of (12). Based on the framework of this principle, the problem of (12) can be iteratively solved by the proposed cross entropy method in the next section. It should be noted that the state equation of the process in (12) is the transition between activity $i-1$ and activity i via route r_i , and (12) is solved in Step 2 of the main algorithm by the agents (described in Section 3.1).

Note that the expectation in (12) cannot be calculated on routes by travelers because its value depends on the net activity values obtained in the future. As a result, the expectation of the travel choice is approximated as the sample average when all agents have completed their activity program. The experienced performance of travel choices influences related choice probabilities on next iteration. The solution of (12) is iteratively approximated within the process of the solution algorithm (equations (19)-(21)).

Given the above optimal decision rule, the activity-chain-based DUE is the optimal solutions of the equation (12). The difficulty remains on how to derive optimal probability distributions towards DUE on a multiagent system.

3. Solution algorithm

To solve the aforementioned DUE problem, a CE-based approach is developed, aiming to derive iteratively optimal probability distributions towards DUE with respect to destination chain, departure time and route choice. The CE approach derives optimal choice probability distributions based on average performance experienced by travelers at previous day (iteration) such that travelers of the same class, competing for the same activity and transport supply, shift to more valuable alternatives at next day. The proposed algorithm is based on traveler's day-to-day learning behaviour, reflecting similar mechanism of traveler's travel choice adjustment process. The experimental investigation of traveler's day-to-day route choice and departure time choice dynamics confirmed that travelers adjust their daily travel decision in response to traffic congestion and the system may converge to user equilibria (Mahmassani 1990). The reader is referred to (Ma and Lebacque 2007; Ma, 2007) for more detail descriptions.

To avoid the route enumeration problem, traveler's route is constructed in a sequential way by selecting an outgoing link at each node until arriving to the activity destination. We have associated with each node a set of OD dependent outgoing arc/node choice probabilities to guide agents towards cheaper routes. The route choice probability is the multiplication of sequential arc choice probabilities of its route. The rationale of this approach is that the sequence of probability distributions should converge towards a distribution concentrated on optimal routes. This emulates the day-to-day learning of agents. The detail of proposed algorithm for solving activity-chain-based DUE problem is described as follows.

3.1 CE approach for the activity-chain-based dynamic traffic assignment problem

The basic concept of the CE approach for dynamic traffic assignment problem is recalled here. Consider a set of routes R_k connecting a pair of origin and destination k . Agents choose a route following a probability distribution \mathbf{p} , which is iteratively adjusted with respect to the performance of the route. Let $H_r(\gamma)$ be the performance function of route r , defined by Boltzmann distribution with the control parameter γ :

$$H_r(\gamma) = e^{-C_r(d_r)/\gamma}, \quad \forall r \in R_k \quad (13)$$

where $C_r(d_r)$ is the travel cost of route r depending on its travel demand d_r . The parameter γ controls the swapping force pushing agents shift from more costly routes to cheaper ones. As its value increases, the shifting force reduces. The overall expected performance based on the choice probability distribution \mathbf{p} is the expectation of all route performance. Following Rubinstein (1999), the optimal probability distribution towards cheaper routes at the next iteration is the solution of the following optimization problem,

which minimize the Kullback-Liebler relative entropy between two consecutive probability distributions:

$$\mathbf{p}^{w+1} = \max_{\mathbf{p}} \mathbb{E}_{\mathbf{p}^w} [H(\gamma) \ln \mathbf{p}] \quad (14)$$

subject to

$$\sum_{r \in R_k} p_r = 1, \quad \forall p_r \geq 0 \quad (15)$$

with w being the index of iteration. The idea here is that \mathbf{p}^{w+1} should be as close as possible to the distribution \mathbf{p}^w weighted by the performance distribution (which favors low cost routes).

In the case of a single OD and many routes, the optimal probability distribution obtained by solving the above optimization problem is given by

$$p_r^{w+1} = p_r^w \frac{e^{-C_r^w / \gamma^w}}{\sum_{s \in R_k} p_s^w e^{-C_s^w / \gamma^w}}, \quad \forall r \in R_k \quad (16)$$

Following previous work (Ma and Lebacque, 2007), the control parameter is determined by solving the following optimization problem:

$$\text{Min } \gamma^w \quad \text{subject to } \sum_{r \in R_k} |p_r^{w+1} - p_r^w| \leq \alpha^w \quad (17)$$

where $\alpha^w = \frac{C}{w}$ is a numerical divergent series such that the flow adjustment converges to fixed points. C is a positive constant for setting initial value of γ^w .

We conjecture that an iterative process such as (17) and (19)-(21) converges towards fixed points. The reasoning is the following:

- imposing (17) implies that the γ^w will (on average) increase,
- then the process of (16) will approximate the integration of field lines of the field obtained by letting $\gamma \rightarrow \infty$ in (16). The reader is referred to the articles of Ma and Lebacque (2007) and Lebacque et al. (2009) for detailed description.
- The integrated lines of the field converge towards the fixed points of the field. Numerical experiments support this conjecture.

The solution algorithm for the multiagent dynamic traffic assignment problem iteratively updates the related choice probability distributions towards the DUE. The procedure of simulation is described as follows. Each agent constructs a route chain connecting activity destinations of his/her activity program. Conditional with the

destination chain determined at the origin, an agent selects his departure time for next activity at current activity location, and then incrementally moves to the next outgoing arc until the destination. When arriving at destination, the agent chooses one vacant activity with best gross activity value. Its net activity value can then be evaluated by reducing the general travel cost of trip from obtained gross activity value. This process is repeated until agent's activity program is accomplished. When all agents complete their activity programs, the probability distributions with respect to destination-chain, departure time and outgoing node choice are updated according to the performance of experienced travel alternatives of agents.

The detail of proposed solution algorithm for activity-chain-based DUE is described as follows.

Main algorithm

Step 1: Initialization

Initialize the uniform probability distribution for destination-chain and departure time choice for all travelers. The interval for departure time choice from activities $i-1$ to i is initialized within the desired starting time $[\tau_{i-1}, \tau_i]$. Let an outgoing node y at x be a successor of x . For outgoing node choice at node x , a set of vectors of probabilities is associated with node x , $\forall x \in V$ for all OD pair $k \in K$. Agents utilize these probability distributions to construct their routes to destinations. The probability vector \mathbf{p}_{xk} (see definition below) is initialized by applying an iteratively updating procedure, suggested by the solution of Eq. (14) (Rubinstein, 1999).

- 1.1 Initialize a uniform distribution for conditional outgoing node choice for all nodes. Let $\mathbf{p}_{xk}^z = \{p_{y|xk}^z | y \in \Lambda^+(x)\}$, $\forall p_{y|xk}^z \geq 0$ denotes the vector of conditional choice probabilities $p_{y|xk}^z$ for outgoing node y at node x for OD pair k at iteration z . $\Lambda^+(x)$ is the set of outgoing nodes of x , $\forall x \in V$ and $\forall k \in K$. Set iteration index $z=1$.
- 1.2 Generate N random routes for all $k \in K$ based on \mathbf{p}_{xk}^z . Evaluate route costs and order them from lowest to highest. Then calculate the ρ -quantile of the route costs $C_{\lceil \rho N \rceil}$, i.e. $\lceil \rho N \rceil$ th lowest cost ($\lceil a \rceil$ denotes the smallest integer greater than or equal to a). Note that ρ is set small, say $0.1 \leq \rho \leq 0.3$, such that resulting outgoing node choice probability is not too small.
- 1.3 Update $p_{y|xk}^z$ as:

$$p_{y|_{xk}}^{z+1} = \frac{\sum_{r \in R_k} 1_{[C_r \leq C_{\lceil \rho N \rceil}]}(x, y)}{\sum_{r \in R_k} 1_{[C_r \leq C_{\lceil \rho N \rceil}]}(x)}, \forall y \in \Lambda^+(x) \quad (18)$$

where $1_{[C_r \leq C_{\lceil \rho N \rceil}]}(x, y)$ is an indicator being 1 if the arc (x, y) belongs to route r and the cost of route r satisfies $C_r \leq C_{\lceil \rho N \rceil}$, and 0 otherwise.

1.4 If $|\mathbf{p}_{xk}^{z+1} - \mathbf{p}_{xk}^z| \leq \boldsymbol{\varepsilon}$, $\boldsymbol{\varepsilon} > \mathbf{0}$, then stop. Otherwise go to 1.2. Set $z = z + 1$.

Note that the above initialization of conditional outgoing node choice probability favors travelers' use of shortest paths between origins and destinations. Also, it should keep the probabilities of other outgoing node choices not too small so that the utilization of other routes is also possible.

Step 2: Route construction and net activity value evaluation

Load agents in the network according to the aforementioned simulation procedure. Each agent constructs its trip chain to complete his/her fixed activity program g based on related choice probability distributions. The departure time choice horizon T_i for the next scheduled activity i is considered between agent's arrival time t_{i-1} at activity $i-1$, and the desired starting time τ_i of next activity i . Agents' choice of departure time and outgoing node are conducted at the activity destinations and at the nodes of network, respectively, according to related probability distributions. When all agents have completed their activity program, the average performance of related destination chain, departure time and route choice can then be evaluated by (5). The approximate solution to the sequential optimisation problem of (12) is iteratively ameliorated by updating the choice probabilities in Sep 3.

Step 3: Choice probability update

The choice probability distributions with respect to destination-chain, departure time and outgoing node choice are updated as follows.

Destination chain choice probability update

The choice probability of destination-chain \mathbf{d} (a sequence of destinations), conditional on agent's activity program g and his/her origin o , is updated based on the summation of normalized net activity values obtained at activity destinations. It is defined as:

$$p_{\mathbf{d}^s|og}^{w+1} = p_{\mathbf{d}^s|og}^w \frac{e^{-\tilde{v}_{\mathbf{d}^s|og}^w / \gamma_{og}^w}}{\sum_{\mathbf{d}^l \in \mathbf{D}_{og}} p_{\mathbf{d}^l|og}^w e^{-\tilde{v}_{\mathbf{d}^l|og}^w / \gamma_{og}^w}}, \quad \forall g, \mathbf{d} \quad (19)$$

with $\tilde{v}_{dog}^w = \frac{v_d^{w,\max} - \hat{v}_{dog}^w}{v_d^{w,\max}}$, where $\hat{v}_{dog}^w = \frac{\bar{v}_{dog}^w}{\bar{v}_{og}^w}$ with \bar{v}_{dog}^w being the average net value of activities for agents of the same destination-chain choice \mathbf{d} , origin o and activity program g . \bar{v}_{og}^w is the average net activity value with respect to origin o and activity program g . The maximum of the net activity value with respect to \mathbf{d} is $v_d^{w,\max} = \max \{ \hat{v}_{dog}^w \mid \forall \mathbf{d} \in \mathbf{D}_{og} \}$ with \mathbf{D}_{og} being the set of possible destination-chains with respect to o and g . γ_{og}^w is the control parameter with respect to o and g at the iteration w , obtained by solving (17). The net activity value obtained by agents at destination is evaluated by (5) and (3).

Departure time choice probability update

The probability of discretized departure time choice h , conditional on agent's activity program g , OD pair k , and desired starting time τ of next activity i , is updated as follows:

$$p_{h^s|kg\tau_i}^{w+1} = p_{h^s|kg\tau_i}^w \frac{e^{-\tilde{v}_{h^s kg\tau_i}^w / \gamma_{kg\tau_i}^w}}{\sum_{h^l \in H_{T_{\tau_i}}} p_{h^l|kg\tau_i}^w e^{-\tilde{v}_{h^l kg\tau_i}^w / \gamma_{kg\tau_i}^w}}, \forall k, g, \tau, h \quad (20)$$

where $\tilde{v}_{hkg\tau_i}^w = \frac{v_{kg\tau_i}^{w,\max} - \hat{v}_{hkg\tau_i}^w}{v_{kg\tau_i}^{w,\max}}$ is the relative normalized net value of activities with

$\hat{v}_{hkg\tau_i}^w = \frac{\bar{v}_{hkg\tau_i}^w}{\bar{v}_{kg\tau_i}^w}$. $v_{kg\tau_i}^{w,\max} = \max \{ \hat{v}_{hkg\tau_i}^w \mid \forall h \in H_{T_{\tau_i}} \}$ denotes the maximum of normalized net value of activities for agents of the same attributes with respect to (k, g, τ_i) . $H_{T_{\tau_i}}$ is the set of the discretized time intervals within the time interval $T_{\tau_i} = [t_{i-1}, \max(t_{i-1}, \tau_i)]$, $\forall i$. Similarly, $\bar{v}_{hkg\tau}^w$ and $\bar{v}_{kg\tau}^w$ is the average net value of activities with respect to (h, k, g, τ) and (h, k, τ) , respectively.

Outgoing node choice probability update

The choice probability of next outgoing node y , conditional on being at node x and agent's OD pair k , is updated as follows:

$$p_{y|xk}^{w+1} = p_{y|xk}^w \frac{e^{-\tilde{c}_{yxx}^w / \gamma_{xk}^w}}{\sum_{y' \in \Lambda^+(x)} p_{y'|xk}^w e^{-\tilde{c}_{y'xk}^w / \gamma_{xk}^w}}, \forall x \in V, \forall k \in K \quad (21)$$

where $\tilde{C}_{yxk}^w = \bar{C}_{yxk}^w / \bar{C}_{xk}^w$ is the normalized average generalized travel cost of OD pair k . \bar{C}_{yxk}^w and \bar{C}_{xk}^w , represent agents' realized average generalized travel cost for routes passing the arc (x, y) and the node x for OD pair k at iteration w , respectively.

Step 4: Stopping criteria

If the derived choice probability distributions for destination chain, departure time and outgoing node choice vary within a small number $\varepsilon > 0$, then stop. Otherwise go to Step 2. Set $w = w + 1$.

Based on the above plausible day-to-day learning process, the approximations of dynamic user equilibrium can be iteratively achieved.

4. A numerical example

This section presents numerical results for a grid road network of 60 nodes and 208 arcs with multiple ODs and multiclass users. The length of directed arcs (road section) is randomly set between 1 km and 5 km. The number of lanes is set uniformly as 2 for all arcs. Two different agents (users) are present on the network: car and bus with different travel speed settings. The travel speed is set as 90 km/hr for car agents and 45km/hr for bus agents. The traffic dynamics modeling is based on the aforementioned extended point queue model. The demand is set as 200 for each class of agents at two origins: nodes 11 and 50 (Fig. 1). Seven activity destinations are set at node 7, 14, 26, 28, 35, 42, and 56. The parameter m_d (average gross activity value at destination d) in Eq. (3) is set respectively as 10, 20, 40, 30, 25, 15, and 55 (euros) for the above destinations. λ_d is set as 4 for all destinations. For simplicity, each agent has an activity program to be implemented, composed of two activities with desired starting time at 9:00 and 11:00, respectively. The activity supply is set as sufficiently large (4000 activities at each destination) for all destinations. The discretized departure time choice interval is 5 minutes. The unitary travel cost and early/late arrival penalty are set as 7 euros/hour, 4 euros/hour and 15 euros/hour, respectively.

[place Fig. 1 about here]

The performance of proposed algorithm is shown in Fig. 2, 3, 4, 5, 6 and 7. Fig. 2 shows the convergence result of total net activity values of agents with respect to origins and transport mode. The difference of total net activity values between 4 classes (2 agent classes multiplied by two origins) reflects that car agents have more advantage to occupy

better activities (lower travel time to arrive to the destinations with higher activity values). Fig. 3 The Wardrop-type dynamic user equilibrium condition can also be verified by agent's obtained total net activity value based on his/her activity program realization. As shown in Fig. 3, most agents of the same class have obtained nearly the same net activity values with only few exceptions. Fig. 4 depicts the convergence of the generalized travel costs for car and bus agents departing from node 11. The ordered travel cost of these agents converges towards a lower value due to the fact that agents learn gradually better departure time and route choices. Similarly, agents learn to shift to more attractive destination chain and departure time choices. Note that the number of car/bus agents between node 26 and node 56 is eventually greater than 200 since it combines agents from two possible origins. Fig. 5 depicts the evolution of choice probability over possible destination chains for bus agents from origin 11. The results show that the probabilities of best destination-chain choice node 26 and node 56 increase rapidly to 1, due to its higher gross activity values (40 and 55 euros per activity, respectively). However, the probabilities of other destination-chain choice alternatives decrease rapidly to 0. Fig. 6 shows the evolution of choice probabilities for the departure time choice. The agents modify gradually the departure time choice with respect to the desired arrival time. The departure time choice probability converges iteratively to some time intervals with higher performance. Fig. 7 shows the evolution of choice probability over outgoing nodes along the shortest route 11-12-13-14-15-16-26. The results illustrate that the choice probability of outgoing nodes on the shortest route increase gradually. Despite the fact that agents reinforce the choice probability over the links on the shortest route, the choice probability over other links have non-zero probabilities due to the fact that the difference of normalized generalized costs is relatively small, influenced jointly by the departure time choice of agents.

[place Fig. 2 - Fig. 7 about here]

5. Conclusion and discussions

In this paper, a dynamical activity-chaining model is proposed to capture individual's travel behavior under traffic dynamics and activity supply. A multiagent approach is implemented to model the individual's travel decision choice process with an activity supply model based on the Accessibility to Vacant Activities model. We propose a cross entropy based solution algorithm to obtain approximate solutions of the multiclass dynamic traffic assignment problem in a multiagent transportation system. Different with route-based flow adjustment algorithms, the proposed method is arc-based, avoiding the route enumeration issue for the dynamic traffic assignment problem. A

significant contribution of this paper is addressing the dynamic user equilibrium issue in a multiagent context. By assuming bounded rationality of user behavior, an individual's best strategy at the activity-chaining process is to maximize the expectation of utility of choice alternatives under uncertainty. Based on the framework of Bellman's principle of optimality, the proposed cross entropy algorithm drives the system towards approximations of dynamic user equilibriums. The numerical study on a bimodal simulation network shows that the algorithm converges to such an approximation of the equilibrium. Although the proposed CE approach is an exact algorithm rather than a heuristic method, the user equilibrium condition is usually difficult to verify in stochastic environment as presented in our numerical study.

Future investigations and extensions include the study of the convergence of the algorithm and also the equilibrium stability in the presence of multiple equilibriums. Moreover, further research is needed concerning its efficiency when the number of OD pairs and the size of the network increase. The comparison of the performance of the proposed algorithm with other algorithms is also necessary, which is currently under study.

Acknowledgements

The authors much appreciate valuable comments and suggestions from three anonymous referees. This research has benefited from a grant of ANR (the French Agency for Research). Project EuroCities-DATTA n° ANR-07-BLAN-0032-01.

Reference

1. Arentze, T., Timmermans, H.J.P., 2004. A learning-based transportation oriented simulation system. *Transportation Research part B* 38(7), 613-633.
2. Axhausen, K.W., 1990. A simultaneous simulation of activity chains and traffic flow. In: Jones, P. (Ed) *Development in Dynamic and Activity-Based Approaches to Travel Analysis*, Avebury, Aldershot, England, pp. 206–225.
3. Bertsekas, D.P., Gafni, E., 1982. Projection methods for variational inequalities with applications to the traffic assignment problem. *Mathematical Programming Studies* 17, 139-159.
4. Buisson, C., Lebacque, J.P., Lesort, J.B., 1995. Macroscopic modelling of traffic flow and assignment in mixed networks. In: Pahl, P.J., Werner, H. (Eds.) *Proceedings of the Berlin ICCCB E Conference* 1367-1374.
5. Busoniu, L., Babuska, R., De Schutter, B., 2008. A comprehensive survey of multiagent reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics Part C* 38(2), 156-172.
6. Cetin, N., Nagel, K., Raney, B., Vollmy, A., 2002. Large-scale multi-agent transportation simulations. *Computer Physics Communications* 147, 559-564

7. Daganzo, C.F., 1994. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research part B* 28(4), 269-287.
8. Dafermos, S.C. 1972. The Traffic Assignment Problem for Multiclass-User Transportation Networks. *Transportation science* 6(1), 73-87.
9. Dafermos, S.C. 1980. Traffic Equilibrium and Variational Inequalities. *Transportation science* 14(1), 42-54.
10. Dafermos S.C. 1982. Relaxation algorithms for the general asymmetric traffic equilibrium problem. *Transportation Science* 16, 231-240
11. Elloumi, E., Haj-Salem, H., Papageorgiou, M. 1994. METACOR, a macroscopic modelling tool for urban corridors. In: *Proceedings of the 2nd Triennial Symposium on Transportation Analysis, Capri*.
12. Fudenberg, D., Levine, D. 1998. *Theory of Learning in Games*. MIT Press.
13. Garavello, M., Piccoli, B., 2006. Traffic flow on road network using the Aw-Rascle model. *Communications partial differential equations* 31, 243-275.
14. Helvik, B.E., Wittner, O., 2001. Using the cross-entropy method to guide/govern mobile agent's route finding in networks. In: Pierre, S., Glitho, R.H. (Eds.) *Proceedings of the Third international Workshop on Mobile Agents For Telecommunication Applications. Lecture Notes In Computer Science*, vol. 2164. Springer-Verlag, London, pp.255-268.
15. Hu, J., Wellman, M.P. 2003. Nash Q-learning for general-sum stochastic games, *Journal of Machine Learning Research* 4, 1039-1069.
16. Jin, W.L., 2007. A dynamical system model of the traffic assignment problem. *Transportation Research part B* 41, 32-48.
17. Jones, P.M., Koppelman, F.S., Orfueil, J.P., 1990. Activity analysis: state-of-the-art and future directions. In: Jones, P.M. (Ed.), *New Developments in Dynamic and Activity-Based Approaches to Travel Analysis*. Gower Publishing: Aldershot, England, pp. 34–55.
18. Khoshyaran, M.M., Lebacque, J.P., 2008. Lagrangian modeling of intersections for the GSOM generic macroscopic traffic flow model. In: *Proceedings of the 10th International Conference on Application of Advanced Technologies in Transportation (AATT2008)*, Athens, Greece.
19. Khoshyaran M.M., Lebacque J.P. 2009 Internal state models for intersections in macroscopic traffic flow models. *Traffic and Granular Flow '09*, Springer Berlin Heidelberg. (to appear)
20. Kuwahara M., Akamatsu, T., 1997. Decomposition of the reactive assignments with queues for many-to-many origin-destination pattern. *Transportation Research part B* 31(1), 1-10.
21. Lam, W.H.K., Yin, Y, 2001. An activity based time-dependent traffic assignment model. *Transportation Research part B* 35, 549-574.

22. Lebacque, J.P., 1996. The Godunov scheme and what it means for first order traffic flow models. In: Lesort, J.B. (Ed.), Proceedings of the 13th International Symposium on Transportation and Traffic Theory, Amsterdam: Pergamon, pp. 647-678.
23. Lebacque J.P., Khoshyaran, M.M., 2005. First order macroscopic traffic flow models: intersection modeling, network modeling. In: Mahmassani, H.S. (Ed.), proceedings of the 16th International Symposium on Transportation and Traffic Theory, Elsevier, pp. 365-386.
24. Lebacque J.P., Mammam S., Haj-Salem H., 2007. Generic second order traffic flow modeling. In: Proceedings of the 17th International Symposium on Transportation and Traffic Flow Theory. London, pp.755-776.
25. Lebacque, J.P., Ma, T.-Y. Khoshyaran, M.M., 2009. The cross-entropy field for multi-modal dynamic assignment. Traffic and Granular Flow '09, Springer Berlin Heidelberg. (to appear)
26. Leurent, F., 1999. Accessibility to Vacant Activities: a Novel Model of Destination Choice. In: Proceedings of the European Transport Conference, Seminar F, PTRC, London, pp. 307-318.
27. Littman, M.L., 1996. Algorithms for sequential decision making. PhD Dissertation. Brown University, USA.
28. Lo, H.K., 1999. A dynamic traffic assignment formulation that encapsulates the Cell transmission model. In: Ceder, A (Ed.), Proceedings of the 14th International Symposium on Transportation and Traffic Theory, Pergamon, pp. 327-350
29. Ma, T.Y., Lebacque, J.P., 2006. Modeling activity choice distribution and road choice behaviour on a network with side constraints. In: Proceedings of the 11th Euro Working Group on Transportation.
30. Ma, T.Y., Lebacque, J.P., 2007. 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, pp. 161-170.
31. Ma, T.Y., 2007 Modèle dynamique de transport basé sur les activités. PhD dissertation. Ecole Nationale des Ponts et Chaussées.
32. Ma, T.Y., Lebacque, J.P., 2008. A dynamic packet-based multi-agent approach for large scale multimodal network simulation. Proceedings of the 10th International Conference on Application of Advanced Technologies in Transportation (AATT08), Athens, Greece.
33. Mahmassani, H.S., Chang, G.-L., 1985. Dynamic aspects of departure time choice behaviour in a commuting system: Theoretical framework and experimental analysis. Transportation Research Record, No. 1037, 88-101.
34. Mahmassani, H.S., 1990. Dynamic models of commuter behavior: Experimental investigation and application to the analysis of planned traffic disruptions. Transportation Research Part A, 24(6), 465-484.

35. Mammari, S., 2006. Développement d'un modèle de simulation macro-microscopique de trafic. PhD Dissertation. CIFRE PHOENIX-INRETS-Evry University (France).
36. Maruyama, T., Sumalee, A., 2007. Efficiency and equity comparison of cordon- and area-based road pricing schemes using a trip-chain equilibrium model. *Transportation Research part A* 41, 655–671.
37. Meschini, L., Gentile, G. and Papola, N., 2007. A Frequency Based Transit Model for Dynamic Traffic Assignment to Multimodal Networks. *Proceedings of 17th International Symposium on Transportation and Traffic Theory*, London.
38. Nagurney, A., 1993. *Network Economics: A Variational Inequality Approach*. Kluwer Academic Publishers, Norwell, MA.
39. Patriksson, M. 1994. *The Traffic Assignment Problem, Models and Methods*. VSP, Utrecht, The Netherlands.
40. Peeta S, Ziliaskopoulos, A., 2001. Foundations of dynamic traffic assignment: The past, the present and the future. *Networks and Spatial Economics* 1(3-4), 233-265.
41. Ramadurai, G., Ukkusuri, S., 2008. Dynamic user equilibrium model for combined activity-travel choices using activity-travel supernetwork representation. *Networks and Spatial Economics*. DOI 10.1007/s11067-008-9078-3
42. Raney, B., Cetin, N., Völlmy, A., Vrtic, M., Axhausen, K., Nagel, K., 2003. An agent-based microsimulation model of Swiss travel: First results. *Networks and Spatial Economics* 3(1), 23–41.
43. Recker, W.W., 2001. A bridge between travel demand modeling and activity-based travel analysis. *Transportation Research* 35B (5), 481–506.
44. Rieser, M., Nagel, K., Beuck, U., Balmer, M., Rügenapp, J., 2007. Agent-oriented coupling of activity-based demand generation with multiagent traffic simulation. *Transportation Research Record* 2021, 10–17.
45. Rubinstein, R.Y., 1999. The cross-entropy method for combinatorial and continuous optimization. *Methodology and Computing in Applied Probability* 1(2), 127-190.
46. Small, K.A., 1982. The scheduling of consumer activities: Work trips. *American Economic Review* 72(3), 467-479.
47. Smith, M.J., 1979. The existence, uniqueness and stability of traffic equilibrium. *Transportation Research part B* 13, 295-304.
48. Smith, M.J., 1993. A new dynamic traffic model and the existence and calculation of dynamic user equilibria on congested capacity-constrained road networks. *Transportation Research part B* 27(1), 49-63.
49. Tong, C.O., Wong, S.C., 2000. A predictive dynamic traffic assignment model in congested capacity-constrained road networks. *Transportation Research part B* 34(8), 625-644.
50. Vickrey, W.S., 1969. Congestion theory and transport investment, *American Economic Review* 59(2), 251-260.

51. Wardrop, J.G. 1952. Some theoretical aspects of road traffic research. In: Proceedings of the Institution of Civil Engineers, Part II 1, 325-362.
52. Wynter, L., 2001. A convergent algorithm for the multimodal traffic equilibrium problem. Research report, vol. 4125, INRIA, France.

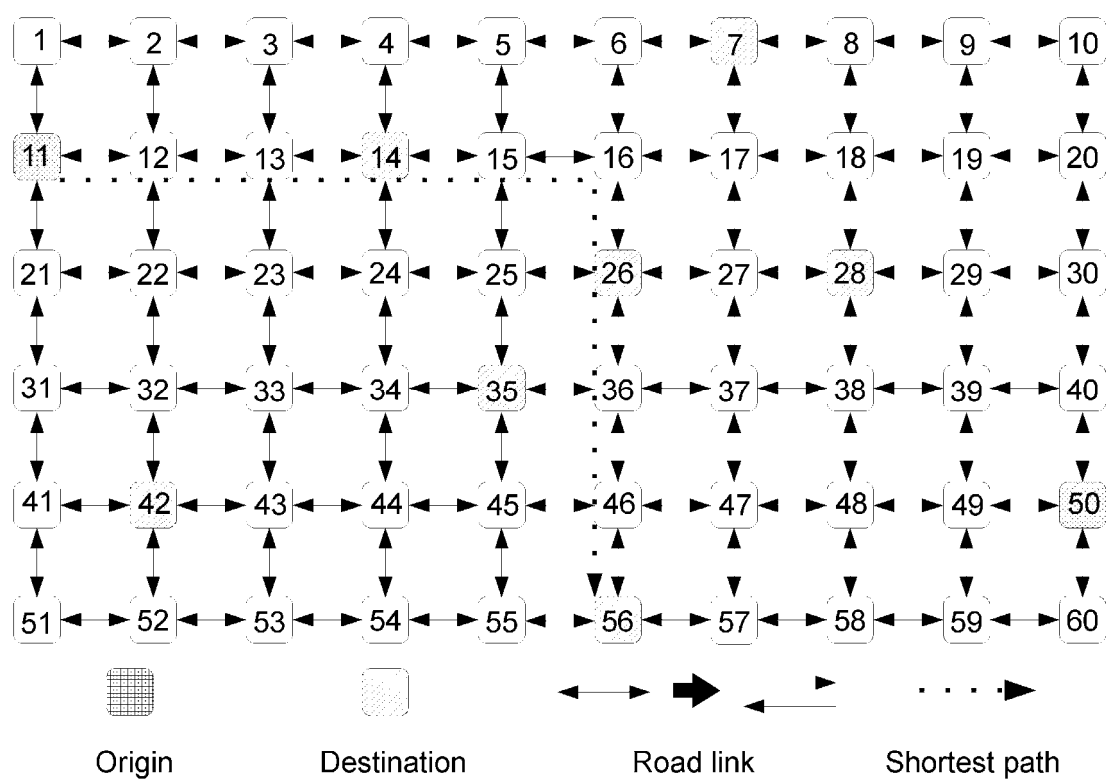


Figure 1. Grid road network example

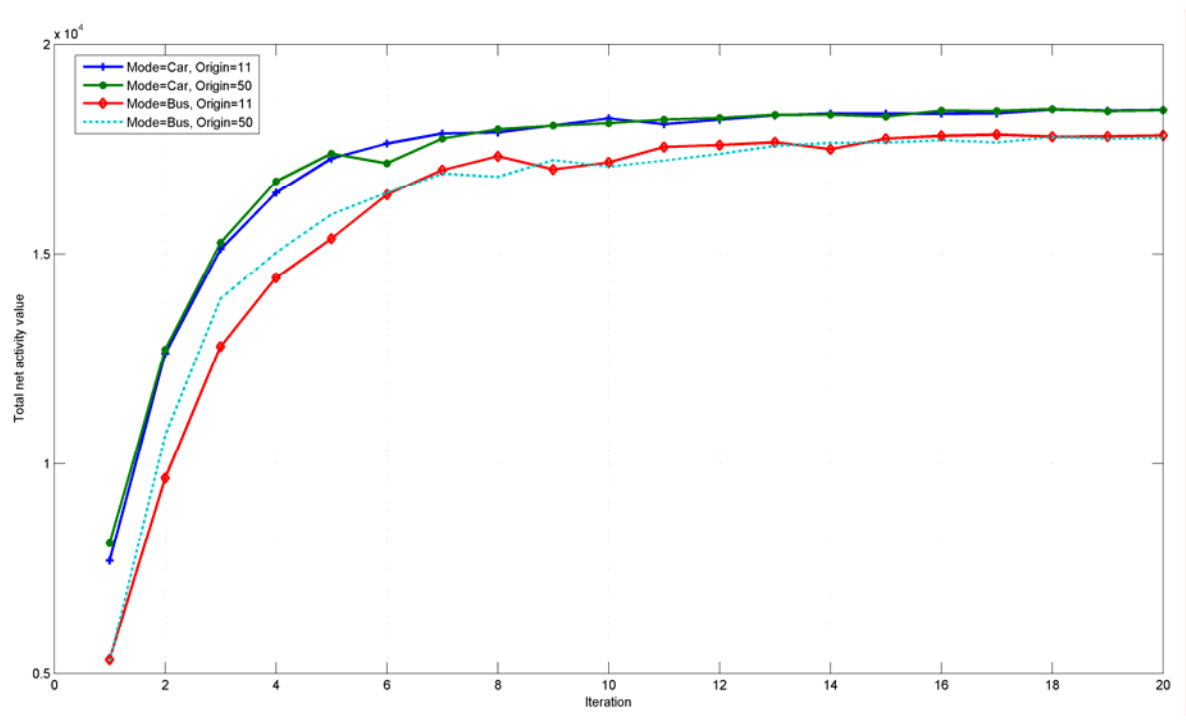


Figure 2. Convergence of all agents' obtained net activity values based on activity program realization

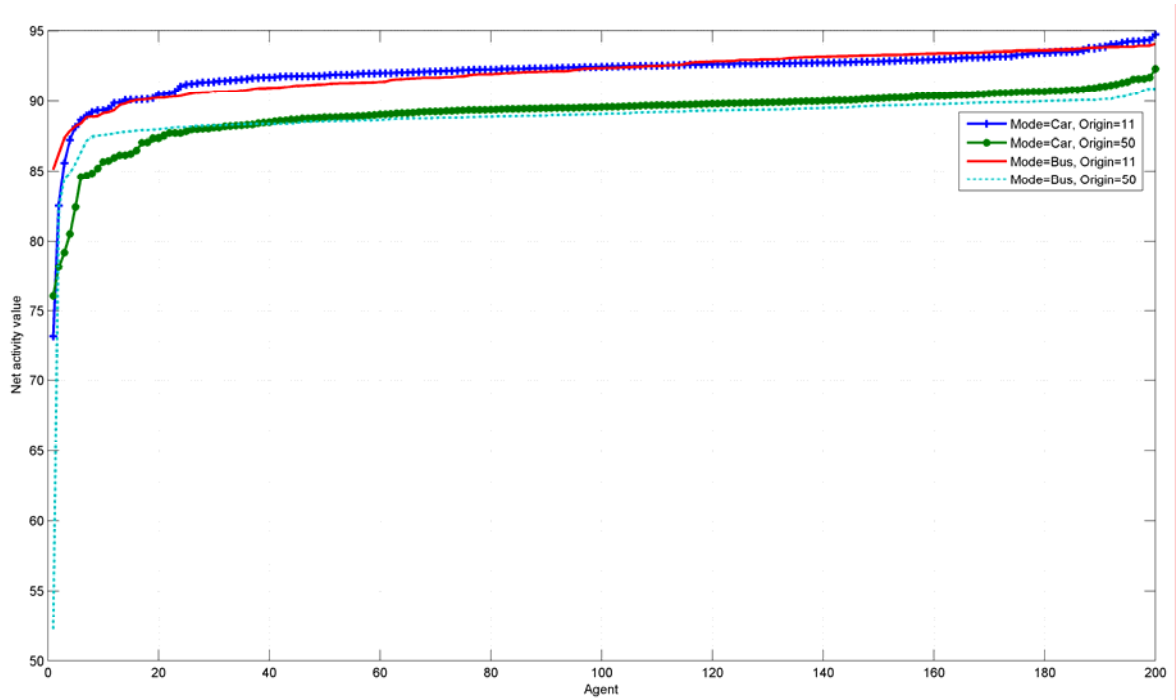


Figure 3. Ordered agent's obtained net activity value, by class and origin, based on activity program realization

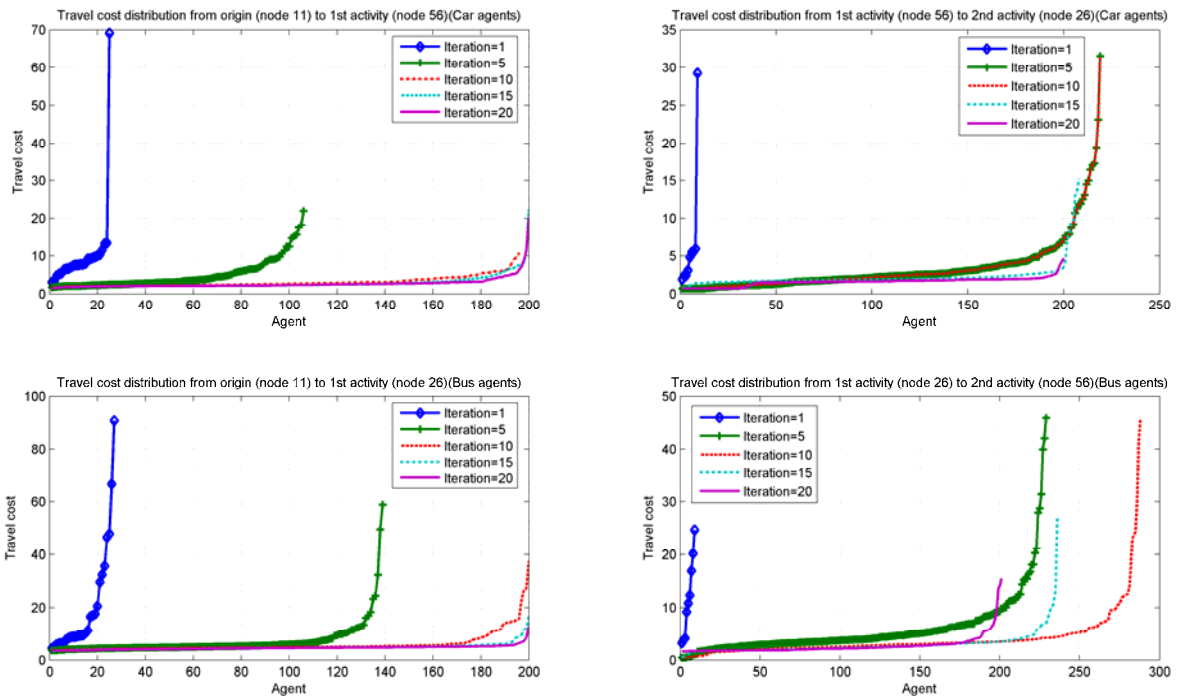


Figure 4. Ordered travel cost of car/bus agents from origin node 11 to the first activity destination (node 56/node 26, respectively), (left), and from node 56/node 26 to the second activity (node 26/node 56, respectively), (right).

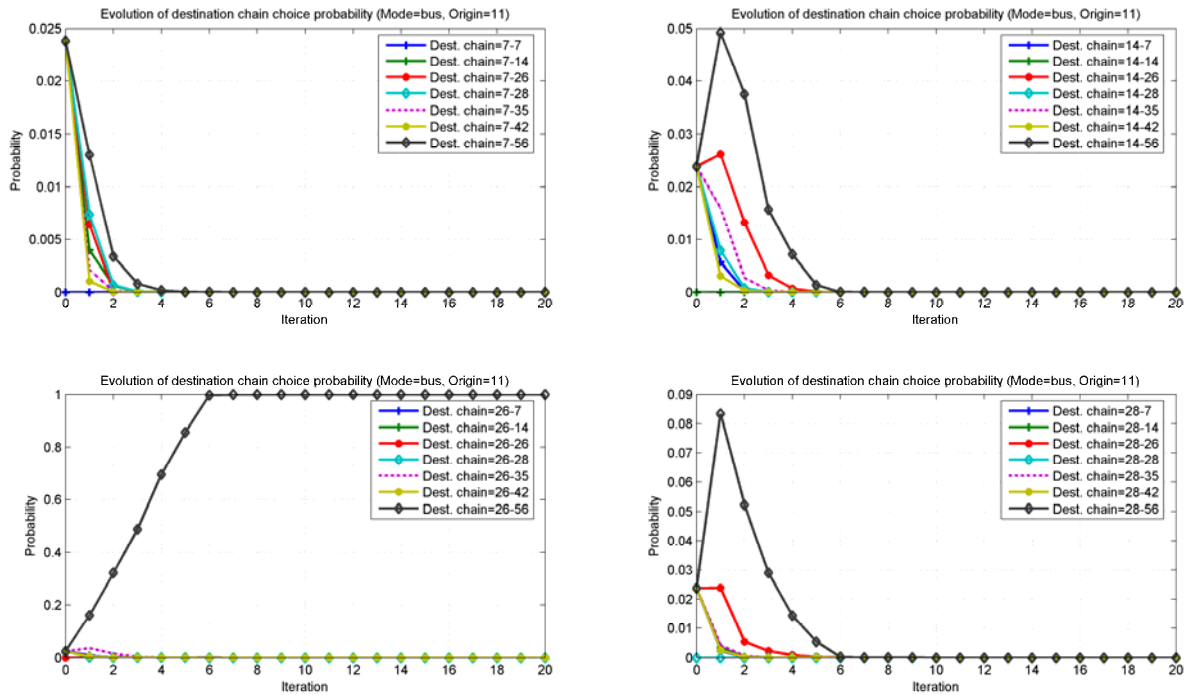


Figure 5. Evolution of destination-chain choice probability for bus agents departing from origin node 11: destination chain beginning at node 7 and 26, (left), destination chain beginning at node 14 and 28, (right).

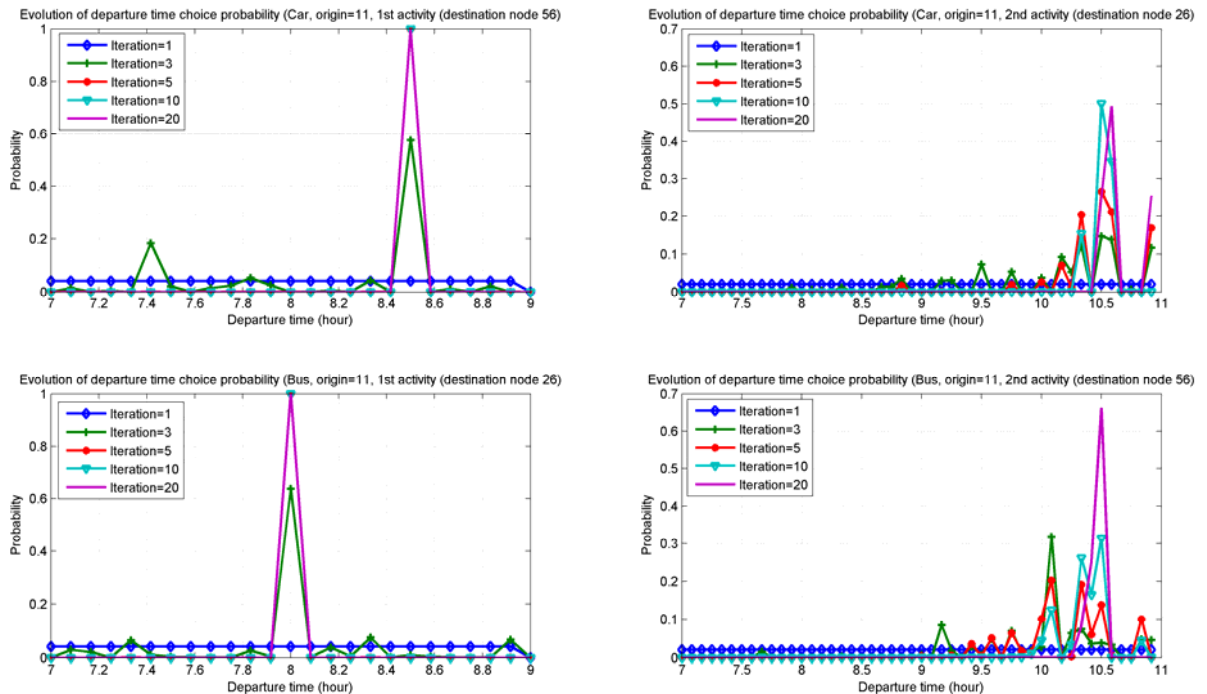


Figure 6. Evolution of departure time choice probabilities for car/bus agents from origin node 11: from node 11 to node 56/node 26, (left), and from node 56/node 26 to node 26/node 56 (right).

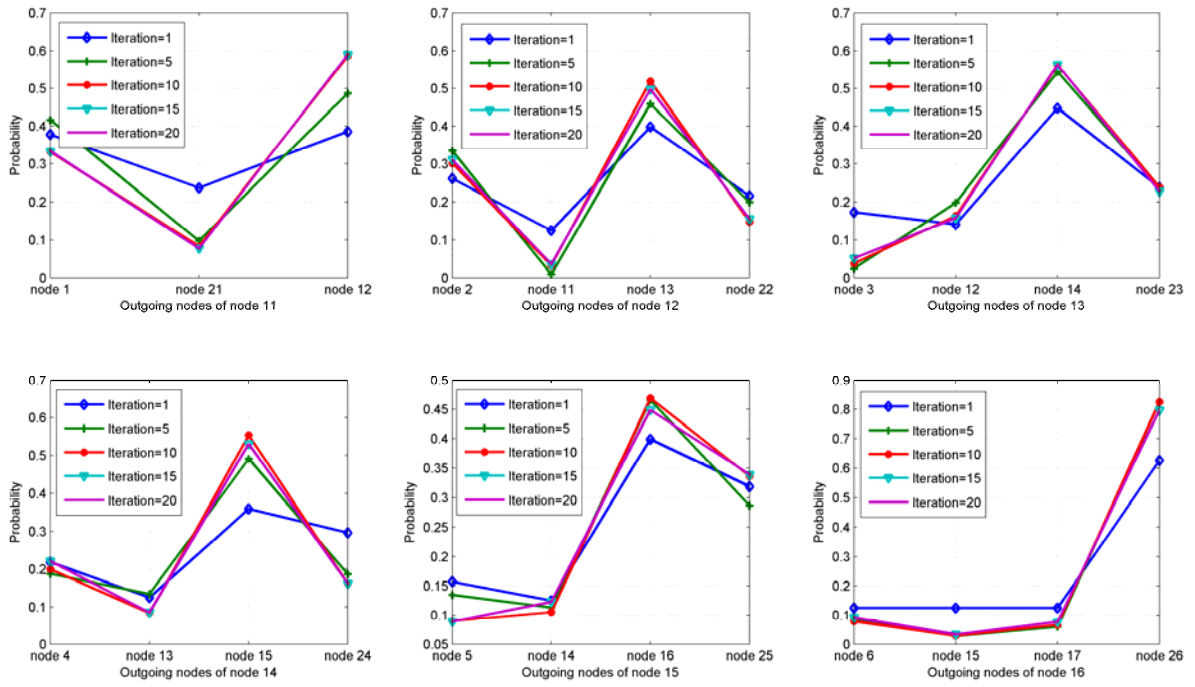


Figure 7. Evolution of outgoing node choice probability associated with nodes on a shortest route 11-12-13-14-15-16-26 for agents from origin node 11 to destination node 26.