Curbing the computational difficulty of the logit equilibrium assignment model

Fabien Leurent

Curbing the computational difficulty of the logit equilibrium assignment model

FABIEN M. LEURENT

INRETS, 2 avenue Malleret-Joinville, 94114 Arcueil Cedex, France

ABSTRACT

In the past, research in traffic assignment modeling has been directed primarily towards improving the deterministic model. Alternative, more behavioral principles were thought to be too demanding computationally.

This paper presents two mathematical contributions that enable one to solve a logit assignment model with flow-dependent travel times at a reduced cost. First, a convergence test for Fisk's minimization program is introduced, based on a duality gap principle. Second, a new definition of Dial's STOCH fixed-time logit assignment procedure is given, in which the set of available paths is defined only once and the computations are re-interpreted.

A numerical experiment indicates that these tools make the logit assignment model very competitive compared to the procedures conventionally used for solving the deterministic model.

KEYWORD

Road Transportation; Traffic Assignment Model; Logit; Optimization

INTRODUCTION

Traffic assignment is the fourth and final step in the conventional travel demand forecasting scheme; by partitioning the origin-destination trip rates between several paths, the assignment program attempts to duplicate the vehicular flows on the network.

Most assignment models assume that trip-makers behave rationally. The most well-known assignment principle is that of Wardrop (1952): that every trip-maker attempts to maximize the utility derived from his transportation choices, in other words to minimize his generalized travel time. Thus, a user-optimal equilibrium is achieved when no traveller may decrease his travel time by unilaterally switching paths.

To account for errors in trip-makers' perception of travel time, Daganzo and Sheffi (1977) defined the stochastic user principle, according to which every trip-maker strives to minimize his/her stochastic generalized travel time. This rule allows for partitioning origin-destination trip-rates between several alternative paths, even if their true travel times differ from each other.

Two stochastic models are of particular interest: the logit model (Dial, 1971) and the probit model (Abraham, 1961; Burrell, 1968; Daganzo and Sheffi, 1977). The latter, though behaviorally more appealing, is hardly practical because only Monte-Carlo procedures are available, unless all paths can be identified. The logit model however, is endowed with both an extremely efficient fixed time assignment procedure (Dial's STOCH2), as well as a convex minimization formulation with a closed-form objective function (Fisk, 1980).

Nevertheless, computational difficulties have prevented the logit model from enjoying more widespread use. Among other drawbacks, Fisk's

objective function was thought difficult to evaluate. Only recently have some solution methods been developed: Chen and Alfa (1991) proposed two heuristic algorithms that can be shown to fail on some examples; Damberg et al. (1993) and Bell et al. (1993) proposed algorithms that look efficient but cannot guarantee a stable set of efficient paths when application data are slightly modified.

In this article, we present two developments which make computation of a logit user equilibrium competitive with its deterministic counterpart. First, we design a theoretically-sound convergence test for an equilibrium algorithm such as the Method of Successive Averages. Second, we modify the definition of the set of available paths in Dial's STOCH2: this procedure is problematic if crudely implemented within an equilibration scheme, as the path set is likely to change from one iteration to the next. We put forward some changes that remedy this flaw.

The organization of the paper is the following: Section 2 states the problem in a formal way. Section 3 introduces the convergence test for Fisk's model. In Section 4, we derive a definition of efficient paths that does not depend on congestion phenomena, inspired by Dial's STOCH2. Furthermore, a related path loading procedure is provided, wherein it is easy to compute all the terms in Fisk's objective function. In Section 5, a numerical experiment is carried out to demonstrate that the Method of Successive Averages, combined with the proposed tools, is indeed a very efficient algorithm when applied to the logit model. Section 6 concludes and suggests some further developments.

2. PROBLEM FORMULATION AND MODELING NEEDS

2.1 Logit equilibrium model

Let *r*-*s* be an origin-destination pair with traffic flow q_{rs}, θ a non-negative parameter, k a path from r to s with deterministic travel time T_{rs}^k and flow f_{rs}^k. In the logit assignment model (Dial, 1971), it is assumed that the path flow f_{rs}^k is proportional to a negative exponential function of the travel time T_{rs}^k:

$$f_{rs}^k = q_{rs} \frac{\exp(-\theta T_{rs}^k)}{\sum_i \exp(-\theta T_{rs}^i)}. \tag{1}$$

Then it is automatically ensured that

$$q_{rs} = \sum_k f_{rs}^k. \tag{2}$$

The travel time of path k is related to the travel times T_a of the links a that belong to it via

$$T_{rs}^k = \sum_{a \in k} T_a = \sum_a \delta_{rs}^{ak} T_a. \tag{3}$$

where $\delta_{rs}^{ak} = 1$ if $a \in k$, or 0 if not.

Let x_a be the traffic flow on link a:

$$x_a = \sum_{rsk} \delta_{rs}^{ak} f_{rs}^k. \tag{4}$$

Let finally t_a be the travel time function of link a (assumed to be continuous and non-decreasing):

$$T_a = t_a(x_a). \tag{5}$$

Then eqns. (1)-(5) define a logit-based equilibrium. Figure 1 illustrates a logit split between two paths. The larger the value of θ, the higher the proportion of trips assigned to the shortest path.

Fig. 1. Proportion of travellers that choose path 1
as a function of θ and the time difference T_2-T_1 (binary case).

2.2 Fisk's minimization program

Fisk (1980) characterized the logit equilibrium with variable travel times as the unique solution to the following convex minimization program (6):

$$\min_{\mathbf{f}} \; J_L(\mathbf{f}) = \sum_a \int_0^{x_a} t_a(x)\,dx + \frac{1}{\theta}\sum_{rsk} f_{rs}^k \log\left(\frac{f_{rs}^k}{q_{rs}}\right) \qquad (6)$$

subject to (2) and (4) and to $f_{rs}^k \geq 0$.

In (6) we replaced Fisk's $\sum_{rsk} f_{rs}^k \log(f_{rs}^k)$ with $\sum_{rsk} f_{rs}^k \log(f_{rs}^k / q_{rs})$ to facilitate the understanding of the relationship between (6) and the computations in the STOCH algorithm. This does not alter the existence and uniqueness results obtained by Fisk.

However, Fisk did not address a crucial question: how should the available paths be defined? In Beckmann's deterministic model (1956), all existing acyclic paths may be considered; but in a logit model a specific definition is required, since the conventional shortest path routines do not automatically find suboptimal paths. In Dial's paper (1971), two alternative definitions of efficient paths are provided, namely STOCH and STOCH2. But these definitions are consistent only with respect to fixed travel times (ie. with constant functions t_a in eqn. (6)), and cannot be used when time is a function of arc flow. A definition of available paths that is consistent with variable travel times will be provided in Section 4. First, we address equilibration issues.

2.3 The Method of Successive Averages

An equilibration algorithm is needed to compute the equilibrium state of a transportation network in which the link travel times depend on traffic flows. One such algorithm, which has been widely used, is the Method of

Successive Averages (MSA). Powell and Sheffi (1982) proved its convergence in applications to minimization programs as Fisk's (provided that the definition of available paths does not vary).

Let us define a Flow-Independent Assignment (FIA) as a path loading procedure that partitions the O-D flow according to the logit rule, based on a given set of available paths. An FIA yields a solution to (6) with constant travel times and a given set of utilized paths.

The MSA equilibration algorithm is comprised of four steps.

Step 0: Initialization.

o Set iteration counter $n := 0$.

o Choose a sequence $\alpha^{(k)}$ of real numbers such that $(0 \le \alpha^{(k)} \le 1)$, $(\Sigma \alpha^{(k)} = +\infty)$ and $(\Sigma (\alpha^{(k)})^2 < +\infty)$.

o Find an initial feasible flow pattern $x_a^{(0)} = x_a(\mathbf{f}^{(0)})$, which may be obtained through an FIA based on free flow link times.

Step 1: Link Travel Time Update.

o Set $t_a^{(n)} := t_a(x_a^{(n)})$.

Step 2: Direction Finding.

o Perform an FIA of traffic of all O-D pairs, based on link travel times $t_a^{(n)}$: this yields a path flow solution $\mathbf{g}^{(n)}$ and an auxiliary arc flow pattern $y_a^{(n)} = x_a(\mathbf{g}^{(n)})$.

Step 3: Link Flow Update.

o Set $x_a^{(n+1)} = x_a(\mathbf{f}^{(n+1)}) := x_a^{(n)} + \alpha^{(n)}(y_a^{(n)} - x_a^{(n)})$.

Step 4: Convergence Test.

o Apply a convergence test: either a maximum number of iterations, or a test on the maximum value (over the arcs a of the network) of the change in $\sum_{k=1}^{n} \alpha^{(k)} x_a^{(k)} / \sum_{k=1}^{n} \alpha^{(k)}$ from the previous iteration n-1 to

the current one n.

If test is satisfied, then terminate; else increment the iteration counter $n := n+1$ and go to step 1.

The MSA has been widely applied to solve Fisk's program. However, the definition of efficient paths has not been adequately addressed. Thomas (1991) wrote that "it seems likely that methods which incorporate definitions of acceptable paths similar to those of Dial and Gunnarsson are intrinsically non-convergent, though in practice users often claim them to be satisfactory in that respect". In the following section, we introduce a theoretically sound convergence test for the equilibration algorithm. Then, in section 4, we provide a formal definition of the efficient paths.

3. A CONVERGENCE TEST FOR THE LOGIT MODEL

We first consider the issue of designing a theoretically sound convergence test for an application of the MSA to Fisk's program. It is based on a duality gap principle inspired from the deterministic model.

3.1 The duality gap principle in the deterministic model

In the deterministic equilibrium assignment model, the objective function reduces to $\mathrm{J}_D(\mathbf{f}) = \sum_a \int_0^{x_a(\mathbf{f})} t_a(x)\,dx$. The usual convergence test is to evaluate a duality gap between the objective function $\mathrm{J}_D(\mathbf{f}^{(n+1)})$ and a lower bound estimate:

$$\mathrm{J}_D(\mathbf{f}^{(n)}) + \nabla \mathrm{J}_D(\mathbf{f}^{(n)}).(\mathbf{g}^{(n)} - \mathbf{f}^{(n)}),$$

where $\mathbf{g}^{(n)}$ is obtained in the Step 2 of the MSA (or equivalently of the Frank-Wolfe's method). Thus, the duality gap is given by:

$$\mathrm{DG}_D^{(n)} = \sum_a t_a^{(n)}(x_a(\mathbf{f}^{(n+1)}) - x_a(\mathbf{g}^{(n)})) = \sum_{rsk} f_{rs}^{k(n+1)}(T_{rs}^{k(n)} - \min_k T_{rs}^{k(n)}).$$

The duality gap $\mathrm{DG}_D^{(n)}$ is always positive, except at equilibrium, at which

point it is zero. Hence, a convergence test involves checking whether DG is close to zero.

3.2 A convergence test for Fisk's model

We suggest applying the duality gap principle to the logit model. Let us denote the entropic part J_E as the difference between the complete logit objective function J_L and the purely deterministic part J_D:

$$\mathrm{J}_E(\mathbf{f}) = \mathrm{J}_L(\mathbf{f}) - \mathrm{J}_D(\mathbf{f}) = \tfrac{1}{\theta}\sum_{rsk} f_{rs}^k \log(\frac{f_{rs}^k}{q_{rs}}). \tag{7}$$

From Fisk's results (1980) applied to (for the moment) constant travel time functions $t_a(x) := t_a(x_a^{(n)})$, the flow vector $\mathbf{g}^{(n)}$ considered in Step 2 of the MSA is the unique solution to the following auxiliary program:

$$\min_{\mathbf{g}} \ \mathrm{J}_{\mathbf{f}^{(n)}}(\mathbf{g}) = \mathrm{J}_D(\mathbf{f}^{(n)}) + \nabla \mathrm{J}_D(\mathbf{f}^{(n)}).(\mathbf{g} - \mathbf{f}^{(n)}) + \mathrm{J}_E(\mathbf{g}). \tag{8}$$

Thus, we have, for all $\mathbf{h} \geq \mathbf{0}$:

$$\mathrm{J}_D(\mathbf{f}^{(n)}) + \nabla \mathrm{J}_D(\mathbf{f}^{(n)}).(\mathbf{h} - \mathbf{f}^{(n)}) + \mathrm{J}_E(\mathbf{h}) \geq \mathrm{J}_D(\mathbf{f}^{(n)}) + \nabla \mathrm{J}_D(\mathbf{f}^{(n)}).(\mathbf{g}^{(n)} - \mathbf{f}^{(n)}) + \mathrm{J}_E(\mathbf{g}^{(n)}).$$

Since J_D is convex, it also holds that:

$$\mathrm{J}_D(\mathbf{h}) \geq \mathrm{J}_D(\mathbf{f}^{(n)}) + \nabla \mathrm{J}_D(\mathbf{f}^{(n)}).(\mathbf{h} - \mathbf{f}^{(n)}).$$

Rearranging terms, we obtain:

$$\mathrm{J}_D(\mathbf{h}) + \mathrm{J}_E(\mathbf{h}) \geq \mathrm{J}_D(\mathbf{f}^{(n)}) + \mathrm{J}_E(\mathbf{g}^{(n)}) + \nabla \mathrm{J}_D(\mathbf{f}^{(n)}).(\mathbf{f}^{(n)} - \mathbf{g}^{(n)}).$$

Hence $\mathrm{LBE}^{(n)} := \mathrm{J}_D(\mathbf{f}^{(n)}) + \mathrm{J}_E(\mathbf{g}^{(n)}) + \nabla \mathrm{J}_D(\mathbf{f}^{(n)}).(\mathbf{f}^{(n)} - \mathbf{g}^{(n)})$ is a lower bound on the optimal value of J.

Then a duality gap associated with the logit objective function is given by $\mathrm{DG}_L^{(n)} = \mathrm{J}_L(\mathbf{f}^{(n+1)}) - \mathrm{LBE}^{(n)}$.

When applying the MSA algorithm to the logit model, it is in general not possible to compute $\mathrm{J}_E(\mathbf{f}^{(n+1)})$, unless all paths are identified. However, for some models, like the one that will be described in Section 4, it is easy to

compute $J_E(\mathbf{g}^{(n)})$. The trick is to evaluate the duality gap with respect to $\mathbf{g}^{(n)}$ and not with respect to $\mathbf{f}^{(n+1)}$. We also suggest the following convergence test, based on functions related to $\mathbf{g}^{(n)}$ rather than to $\mathbf{f}^{(n+1)}$:

If $J_L(\mathbf{g}^{(n)}) - \mathrm{LBE}^{(n)} \le \varepsilon\left(\left|J_L(\mathbf{g}^{(n)})\right| + \left|\mathrm{LBE}^{(n)}\right|\right)$, then terminate and let $\mathbf{g}^{(n)}$ be the solution to the minimization program (6), else return to Step 1.

If true, the test gives a vector that solves the minimization program, by definition of a minimum. Conversely, if the path flow vector **f*** solves the program, then auxiliary vector **g*** that corresponds to **f*** is in fact equal to it and thus the convergence test is satisfied:

from the definition of **f***: $(\nabla J_D(\mathbf{f}^*) + \nabla J_E(\mathbf{f}^*)).(\mathbf{g}^* - \mathbf{f}^*) \ge 0$,

from the definition of **g***: $(\nabla J_D(\mathbf{f}^*) + \nabla J_E(\mathbf{g}^*)).(\mathbf{f}^* - \mathbf{g}^*) \ge 0$,

adding up yields that: $(\nabla J_D(\mathbf{f}^*) - \nabla J_E(\mathbf{g}^*)).(\mathbf{g}^* - \mathbf{f}^*) \ge 0$.

which gives that $\mathbf{g}^* = \mathbf{f}^*$, provided that J_E is strictly convex: this is guaranted for all O-D pairs on which there are two paths or more ; if there is only one path, the assignment is trivial…

Note that if only a relative measure $J_L - \mathrm{LBE}$ of the duality gap is needed, then it is not necessary to compute J_E: the test reduces to a check whether $J_L(\mathbf{g}^{(n)}) - \mathrm{LBE}^{(n)} \le \varepsilon$, in other words whether
$J_D(\mathbf{g}^{(n)}) - J_D(\mathbf{f}^{(n)}) - \nabla J_D(\mathbf{f}^{(n)}).(\mathbf{g}^{(n)} - \mathbf{f}^{(n)}) \le \varepsilon$.

4. DEVELOPMENT OF THE STOCH3 PROCEDURE

The results obtained so far apply to any set of utilized paths under the sole constraint that no path may include more than once a given node. We now define a set of efficient paths that enables one to benefit from the efficiency of Dial's STOCH2.

Most previous logit assignment models have used Dial's second definition of efficient paths, according to which "a path is efficient (reasonable) if every link in it has its initial node closer to the origin than is its final node". The word "closer" refers to the travel time measured from the origin with respect to a current travel time vector that may change from one iteration to the next. Therefore *there was no use trying to compute an objective function* for the logit assignment model!

Three problems had to be tackled:

- to restrict Dial's set of efficient paths so as to limit its size and for each reasonable path not to be much longer than the shortest one.

- to stabilize the definition of efficient paths so that it depends neither on congestion nor on the iteration number; and

- to find a way to compute the entropic part of the objective function, so as to measure the convergence rate.

Subsection 4.1 deals with the first two questions, based on previous work by Tobin (1977) as regards the first question. Subsection 4.2 introduces the STOCH3 procedure, which offers a practical way to perform a fixed-time logit assignment using the efficient paths defined in subsection 4.1. Subsection 4.3 describes a way to evaluate $J_E(\mathbf{g})$ in the STOCH3 model.

4.1 Definition of a stable set of efficient, not-too-long paths

A path is called "STOCH3-efficient" (or reasonable, or available) if it does not include the same node more than once, if every link has its initial node closer to the origin than its final node, and if every link is "reasonable enough" compared to a reference shortest path.

More precisely, let:

- T_a^0 be a reference generalized travel cost for link a; it can stand for the distance, or for the free-flow travel time.

- $C_r^0(n)$ be a reference shortest generalized travel cost from origin r to node n, based on the link costs T_a^0.

- h_r^a be a maximum "elongation ratio" for link a with respect to the origin r; from previous surveys, it may be set to $h_r^a := 1.6$ in interurban studies (USAP, 1992) or $h_r^a \in [1.3; 1.5]$ in urban studies (Tagliacozzo and Pirzio, 1973).

- B_a, E_a be respectively the beginning and end nodes of link a.

Definition 1: a path k from origin r to destination s is STOCH3-efficient iff

(i) it does not comprise more than once a given node;

(ii) $C_r^0(E_a) > C_r^0(B_a) \quad \forall a \in k$;

(iii) $(1+h_r^a)(C_r^0(E_a) - C_r^0(B_a)) \geq T_a^0$, with $h_r^a \geq 0$, $\forall a \in k$.

A link a that satisfies the last two conditions is called STOCH3-reasonable with respect to origin r.

The last condition in Def. 1 limits the number of efficient paths by limiting their total reference generalized travel cost: defining $H_r = \max_a h_r^a$, summing over all links a that are incident to an efficient path k yields that:

$$Length(k) = \sum_{a \in k} T_a^0 \leq (1+H_r)(C_r^0(s) - C_r^0(r)) = (1+H_r) \min_{k'} Length(k')$$

Conversely, if k satisfies $Length(k) \leq (1+H_r) \min_{k'} Length(k')$, it may not be efficient since the two first conditions must hold as well.

Part (ii) of Def. 1 is inspired from Dial's specification STOCH2 (Dial, 1971), and part (iii) from Tobin (1977). Our own contribution is to impose fixed reference travel costs, T_a^0, thus ensuring a stable definition of the efficient paths whatever the congestion phenomena and the current travel costs, T_a.

4.2 The STOCH3 procedure

In the STOCH3 procedure it is necessary to consider, on the one hand, the

reference generalized travel costs to (implicitly) enumerate the available paths, and, on the other hand, the "actual" travel times according to which the O-D flows are partitioned between the paths.

Program variables

n node with reference travel cost $C_r^0(n)$ from origin r .

$O_r(i)$ the i-th node in the order of increasing access cost $C_r^0(n)$ from r.

Ω_r^a indicator variable = 1 if link a is reasonable from r and 0 otherwise.

T_a current travel cost on link a.

$A(a)$ impedance of link a.

$W_A(a)$ link weight that accounts for the importance of a in contributing to a reasonable path.

$W_N(n)$ node weight.

$X_A(a)$ flow on link a from the current origin r.

$X_N(n)$ flow passing through node n from the current origin r.

$F(a)$ total current flow on link a (over all origins).

$Q_r(s)$ the O-D trip rate from origin node r to destination node s.

Index r can be omitted on the variables A, W_A, W_N, X_A and X_N, as they need not be stored after dealing with origin r.

Algorithm STOCH3

Step 0. Overall preliminaries: calculation of reasonable paths.

o From every origin node r, compute the shortest paths to all nodes n, based on the reference link travel costs T_a^0, yielding the reference access costs $C_r^0(n)$ and a labeling $O_r(i)$ of the nodes n in the order of increasing access cost from r. For each link a, set $\Omega_r^a := 1$ if

$(1+h_r^a)(C_r^0(E_a)-C_r^0(B_a)) \geq T_a^0 > 0,\ \Omega_r^a := 0$ otherwise.

Storing the $C_r^0(n)$ and $O_r(i)$ variables enables us to store the reasonable paths at a reduced cost since only node-based variables are involved.

Step 1. Preliminaries for a standard iteration.

o Initialize the total link flow variables *F(a)* to 0.

o Set the link impedances $A(a) := \exp(-\theta T_a)$.

Steps 2, 3 and 4 are to be run for each origin node *r*.

Step 2. Forward pass.

o Set all $W_A(a)$ and $W_N(n)$ to 0. Set $W_N(r) := 1$.

o For each node n taken in the order of increasing reference cost $C_r^0(n)$ (the i-th node to be considered is indicated by $O_r(i)$), for each link a with beginning node $B_a = n$, if $\Omega_r^a = 1$ then compute $W_A(a) := A(a)W_N(n)$ and add $W_A(a)$ to $W_N(E_a)$, else do nothing.

Step 3. Backward pass.

o For each node n, set $\square$ if n is a destination node, 0 otherwise.

o For each node n taken in the order of decreasing reference cost $C_r^0(n)$ (use the labeling $O_r(i)$ in decreasing order), for each link a with end node $E_a = n$, if $\Omega_r^a = 1$ then compute $X_A(a) := X_N(n)W_A(a)/W_N(E_a)$ and add $X_A(a)$ to $\square$, else set $X_A(a) := 0$.

Step 4. Contribution to total link-flows.

o $\forall a,\ F(a) := F(a) + X_A(a)$

At the end of the procedure, the vector **F** gives the flow-independent logit assignment based on link travel times T_a.

A detailed interpretation of the computations in STOCH3 is included in the Appendix.

4.3 Computation of the entropic part of the objective function in the STOCH3 model

It is shown in the Appendix (Lemma 3, Eqn. (17)) that at the end of the forward pass from origin r, it holds that: $W_N(s) = \sum_k \exp(-\theta T_{rs}^k)$. Hence the composite travel cost from r to s defined by Williams (1977) as $S_{rs}(\mathbf{T}) = -\frac{1}{\theta}\log(\sum_k \exp(-\theta T_{rs}^k))$ can be evaluated as

$$S_{rs}(\mathbf{T}) = -\frac{1}{\theta}\log(W_N(s)). \tag{9}$$

Eqn. (9) is very useful since it enables one to evaluate the level-of-service, hence the entropic part of the objective function, without enumerating paths.

As $g_{rs}^k = q_{rs}\exp(-\theta T_{rs}^k)/\sum_k \exp(-\theta T_{rs}^k)$, we get:

$$\frac{1}{\theta}\sum_k g_{rs}^k \log\left(\frac{g_{rs}^k}{q_{rs}}\right) = -\sum_k g_{rs}^k (T_{rs}^k - S_{rs}(\mathbf{T}))$$

$$= -\left(\sum_k g_{rs}^k\ T_{rs}^k\right) + q_{rs} S_{rs}(\mathbf{T})$$

and by summing over all origin-destination pairs r-s,

$$\mathrm{J}_E(\mathbf{g}) = \frac{1}{\theta}\sum_{rsk} g_{rs}^k \log\left(\frac{g_{rs}^k}{q_{rs}}\right) = -\left(\sum_a x_a(\mathbf{g})\, T_a\right) + \sum_{rs} q_{rs} S_{rs}(\mathbf{T}). \tag{10}$$

Then we can evaluate the entropic part of the objective function and apply to the STOCH3 model the convergence test designed in Section 3.

5. COMPUTATIONAL EVIDENCE

In this section, we carry out a numerical example to compare the performance of the STOCH3 logit model using the MSA, with that of the deterministic model using both the Frank-Wolfe algorithm and the MSA.

5.1 The case study

The application is related to the western part of the Paris metropolitan area during the evening peak period, with a typical trip travel time of one hour. The test network is composed of 2,000 directed links. There are 141 origin and destination zones.

The dispersion parameter θ is set to 0.233 mn^{-1}, so that when two routes compete with each other, the first one with a travel time five minutes shorter than the second one, approximately three out of four drivers choose the first road. The elongation ratios h_r^a are set to +•, so as to ensure a comparability with previous applications of Dial's logit model.

5.2 Results

Three assignments were performed:

- a logit assignment using the MSA with step size $\alpha^{(n)} = \frac{1}{4+n/10}$; it is further referred to as MSA-L. The step-size was chosen so as to yield large, though limited, moves as, somehow, in solving a fixed-point problem by the substitution method.

- a deterministic assignment (see sub-section 3.1) using the Frank-Wolfe algorithm, further referred to as FW.

- a deterministic assignment using the MSA with step size $\alpha^{(n)} = \frac{1}{1+n}$, further referred to as MSA-D. In the deterministic case, this step-sizing rule was found more efficient than the one used in the logit case.

The performance of each of these procedures was evaluated by measuring a convergence indicator as $\log\left|\frac{X^{(n)}}{J^*} - 1\right|$, where:

- J^* is the optimal value of the objective function of the model (evaluated after convergence of the algorithm),

- $X^{(n)}$ is the value of the objective function at the n-th iteration. In the case of the logit model, $X^{(n)} = J_L(\mathbf{g}^{(n)})$. In the case of the deterministic model, $X^{(n)} = J_D(\mathbf{f}^{(n+1)})$. Although not representative of a practical application where J^* would not be known a priori, this convergence measure ensures that the testbed is fair to all competing algorithms.

Fig. 2. Convergence rates of the three algorithms.

Figure 2 shows that the convergence rate is much better in the case of the logit model, notably because the descent direction includes information about all of the available paths, not only about the shortest path in each iteration.

6. COMMENTS AND CONCLUSIONS

6.1 Model extensions

In other papers (Leurent 1993a, 1994), I addressed the case of elastic demand and capacity constraints. A dual characterization is available for the capacitated logit equilibrium; it can be formulated as an unconstrained concave maximisation program. The dual program can be solved using Polyak's sub-gradient algorithm (Polyak, 1967; Fukushima, 1984). But, since of the remarkable efficiency of the MSA-L, applying an augmented Lagrangian scheme to the capacitated, primal program is a better option.

The computational efficiency of the MSA applied to the logit assignment model facilitates the following possible extensions of the model:

- diagonalization schemes, for example with travel time functions that depend on flows of several links. It is easy to derive a variational inequality formulation.

- simultaneous models that capture more than one step in the conventional transportation planning process.

- multiple user classes modeling. In an IVHS context, the logit model may be of particular interest for assessing the level of information provided to motorists by a route guidance system (Van Vliet et al. 1990). One way to evaluate the effects of such a system is to model two classes of motorists, the first equipped with a route guidance device and characterized by a large dispersion parameter θ, and the other class of non-equipped drivers characterized by a small θ.

6.2 About path identification

It is computationally useful to identify paths. The STOCH procedure is a way to consider all available paths at a reduced cost. We believe that our numerical experiment demonstrates, above all, that path-based equilibration algorithms can be much more efficient than link-based algorithms, with respect to the number of iterations in the equilibration process. This conclusion is also supported by recent work by Schittenhelm (1990) and Larsson and Patriksson (1992), among others.

Algorithms that identify paths lead to better behavioral models. In the STOCH path loading procedure, the origin-destination flow is partitioned between the paths according to a behavioral rule. Other available behavioral rules are the probit model (Abraham, 1961; Burrell, 1968; see Daganzo and Sheffi, 1977, and Powell and Sheffi, 1982, for a mathematical foundation), and the dual criteria, cost-versus-time model (Marche, 1973; see Leurent, 1993b, for a mathematical foundation). All of these behavioural rules can be efficiently addressed by using algorithms that identify the

paths, either implicitly (Maher and Hughes, 1995, for the probit model) or not (Leurent, 1995, for the dual criteria model). It is thus remarkable that, by the identification of paths, while improving the behavioral aspects of the model, the computational process is greatly facilitated.

ACKNOWLEDGMENTS

The editorial suggestions from Laura MARMORSTEIN and the notes of two anonymous referees are greatly appreciated.

REFERENCES

Abraham, C. (1961). La répartition du trafic entre itinéraires concurrents. *Revue Générale des Routes et Aérodromes*, Paris, June.

Beckmann, M., McGuire, C.B., & Winsten, C.B. (1956). *Studies in the Economics of Transportation*. Yale University Press, New Haven, Connecticut.

Bell, M.G.H., Lam, W.H.K., Ploss, G., & Inaudi, D. (1993). *Stochastic User Equilibrium Assignment and Iterative Balancing*. In C.F. Daganzo (ed.), *Transportation and Traffic Theory*, 427-439. Elsevier, Amsterdam.

Burrell, J.E. (1968). *Multiple Route Assignment and its Application to Capacity Restraint*. 4th International Symposium Theory of Trafffic Flow, Karlsruhe.

Chen, M. , & Alfa, A.S. (1991). Algorithms for solving Fisk's stochastic traffic assignment model. *Transportation Research B*, **25**(6), 405-412.

Daganzo, C.F., & Sheffi, Y. (1977). On Stochastic Models of Traffic Assignment. *Transportation Science*, **11**(3), 253-274.

Damberg, O., Lundgren, J.T., & Patriksson, M. (1993). *An Algorithm for the*

Stochastic User Equilibrium Problem. Proceedings of the 1st meeting of the IFORS-EURO Working Group on Urban Traffic and Transportation, Technical University Munich.

Dial, R.B. (1971). A Probabilistic Multipath Traffic Assignment Model which Obviates Path Enumeration. *Transportation Research*, **5**, 83-111.

Fisk, C. (1980). Some developments in equilibrium traffic assignment methodology. *Transportation Research B*, **14**, 243-256.

Fukushima, M. (1984). On the Dual Approach to the Traffic Assignment Problem. *Transportation Research B*, **18**(3), 235-245.

Larsson, T., & Patriksson, M. (1992). Simplicial Decomposition with Disaggregate Representation for the Traffic Assignement Problem. *Transportation Science*, **26**, 4-17.

Leurent, F. (1993a). *NT 93.6 : Modèle logit à demande élastique et à contraintes de capacité explicites*. Working paper, INRETS, Arcueil, France.

Leurent, F. (1993b). Cost versus Time Equilibrium over a Network. *Eur. J. of Oper. Res.*, **71**(2), 205-221.

Leurent, F. (1994). Logit-based equilibrium assignment model with elastic demand, capacity constraints and a dual solution algorithm. In *Actes INRETS,* **38**, 125-162. INRETS, Arcueil.

Leurent, F. (1995). The practice of a dual criteria assignment model with continuously distributed values-of-time. *23rd European Transport Forum, Proceedings of Seminar E*, 117-128. PTRC, London.

Maher, M.J., & Hughes, P.C. (1995). The attractions of SUE - a practicable approach to stochastic user equilibrium assignment. *23rd European Transport Forum, Proceedings of Seminar E*, 103-115. PTRC, London.

Marche, R. (1973). *Valuation of travel time, French inter-city values*. Paper

presented at the Symposium on Transportation and Environment, Southampton, April.

Polyak, B.T. (1967). A general method for solving extremal problems. *Soviet Math. Doklady*, **8**, 593-597.

Powell, W.B., & Sheffi, Y. (1982). The Convergence of Equilibrium Algorithms with Predetermined Step Sizes. *Transportation Science*, **16**(1), 45-55.

Schittenhelm, H. (1990). On the integration of an effective assignment algorithm with path and path-flow management in a combined trip distribution and traffic assignment algorithm. *18th European Transport Forum, Proceedings of Seminar E*, 203-214. PTRC, London.

Tagliacozzo, F., & Pirzio, F. (1973). Assignment models and urban path selection criteria: results of a survey of the behaviour of road users. *Transportation Research*, **7**, 313-329.

Thomas, R. (1991). *Traffic Assignment Techniques*. Avebury Technical, Aldershot, England.

Tobin, R.L. (1977). An extension of Dial's algorithm utilizing a model of tripmakers' perceptions. *Transportation Research*, **11**(5), 337-342.

USAP (1992). *Comportement des automobilistes en matière de choix d'itinéraires*. Union des Sociétés d'Autoroutes à Péage, Paris, June.

D. Van Vliet (1990). *Modelling route guidance systems using SATURN*, in DRIVE Project V1054 (ASTERIX) Deliverable 4, **B**, 83-105.

Williams, H.C.W.L. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A*, **9**, 285-344.

APPENDIX A. AN INTERPRETATION OF STOCH3'S COMPUTATIONS

We hereafter present some lemmas to demonstrate that STOCH3 computes an elastic demand, logit-based path flow pattern based on the current link travel time vector, as well as to show the purpose of the computations.

We denote as $L_r(n_1,n_2)$ the set of "reasonable paths from node n_1 to node n_2 with respect to origin r", which are the paths from n_1 to n_2 that are only comprised of links a such that $\Omega_r^a = 1$. Of course $L_r(r,s)$ is the set of the STOCH3-efficient paths from r to s.

We note $L_r(n_1,n_2,a) = \{k \in L_r(n_1,n_2)\ ;\ a \in k\}$.

We assume that the parameters θ_{rs} are the same whatever the destination s, from a given origin r. In addition, when dealing with that origin, we can drop the subscript r and only refer to θ.

Let us introduce two shorthand notations, Γ_N and Γ_A, where:

$$\Gamma_N(n_1,n_2) = \sum_{k \in L_r(n_1,n_2)} \exp(-\theta T^k_{n_1 n_2}) \tag{11}$$

$$\Gamma_A(n_1,n_2,a) = \sum_{k \in L_r(n_1,n_2)} \delta^{ak}_{n_1 n_2} \exp(-\theta T^k_{n_1 n_2}) \tag{12}$$

Lemma 1. If $L_r(n_1,n_2,a)$ is non-empty, then there is a 1-1 correspondence between it and the Cartesian product $L_r(n_1,B_a) \times L_r(E_a,n_2)$. Whatever the case, it holds that:

$$\Gamma_A(n_1,n_2,a) = \Gamma_N(n_1,B_a)\,\Gamma_A(B_a,n_2,a) \tag{13}$$

$$= \Gamma_A(n_1,E_a,a)\,\Gamma_N(E_a,n_2) \tag{14}$$

Proof of Lemma 1.

A reasonable path $P = \{a_1,a_2,\ldots a_m\} \in L_r(n_1,n_2,a)$ is such that $\exists i \in \{1,2,\ldots m\}\ \ a_i = a,$ and that $\forall j \in \{1,2,\ldots m\},\ \Omega_r^{a_j} = 1$: it can thus be

decomposed into three subpaths $P_1 = \{a_1, \ldots a_{i-1}\}$, $P_2 = \{a_i\}$, $P_3 = \{a_{i+1}, \ldots a_m\}$ such that $P_1 \in L_r(n_1, B_a)$ and $P_3 \in L_r(E_a, n_2)$. P_1 and P_3 are uniquely determined since a cannot be comprised more than once in P.

Conversely, if $L_r(n_1, n_2, a) \neq \varnothing$, then $\Omega_r^a = 1$: therefore, for every subpath pair $(P_1, P_3) \in L_r(n_1, B_a) \times L_r(E_a, n_2)$, path $P = P_1 + \{a\} + P_3 \in L_r(n_1, n_2, a)$.

To prove (13) and (14), let us first deal with the case $\Omega_r^a = 0$: in that case

$\Gamma_A(n_1, n_2, a) = 0 = \Gamma_A(B_a, n_2, a) = \Gamma_A(n_1, E_a, a)$.

In the case where $\Omega_r^a = 1$, it holds that:

$\delta_{n_1 n_2}^{ak} \exp(-\theta T_{n_1 n_2}^k)$

$= \delta_{n_1 n_2}^{ak} \exp(-\theta T_{n_1 B_a}^{k_1}) \exp(-\theta T_a) \exp(-\theta T_{E_a n_2}^{k_3})$, in which, if $\delta_{n_1 n_2}^{ak} = 1$,

$path_k = path_{k1} + \{a\} + path_{k3} = path_{k1} + path_{k2} = path_{k4} + path_{k3}$

$= \exp(-\theta T_{n_1 B_a}^{k_1}) \delta_{B_a n_2}^{a k_2} \exp(-\theta T_{B_a n_2}^{k_2})$ leading to (13) by summing over indices k_1, k_2 on the right hand side and k on the left hand side, from the 1-1 correspondence between $L_r(n_1, n_2, a)$ and $L_r(n_1, B_a) \times L_r(B_a, n_2, a)$, for $L_r(B_a, n_2, a)$ is isomorphic to $L_r(E_a, n_2)$ if $\Omega_r^a = 1$;

$= \exp(-\theta T_{n_1 E_a}^{k_4}) \delta_{n_1 E_a}^{a k_4} \exp(-\theta T_{E_a n_2}^{k_3})$ leading to (14) by summing over indices k_4, k_3 on the right hand side and k on the left hand side, from the 1-1 correspondence between $L_r(n_1, n_2, a)$ and $L_r(n_1, E_a, a) \times L_r(E_a, n_2)$, for $L_r(n_1, E_a, a)$ is isomorphic to $L_r(n_1, B_a)$ if $\Omega_r^a = 1$.

Lemma 2.

$$W_A(a) = \Gamma_A(r, E_a, a) = \sum_{k \in L_r(r, E_a)} \delta_{r E_a}^{ak} \exp(-\theta T_{r E_a}^k) \tag{15}$$

$$W_N(n) = \Gamma_N(r, n) = \sum_{k \in L_r(r, n)} \exp(-\theta T_{rn}^k) \quad \text{if} \quad n \neq r \tag{16}$$

Proof of Lemma 2.

(16) results straightforwardly from (15) and the computation of $W_N(n)$ in the forward pass: a reasonable path passing through $n \bullet r$ passes through a unique link a with end node n.

We prove (15) recursively, considering the links a in the order of increasing $C_r^0(B_a)$. First we deal with links a such that $\Omega_r^a = 0$: in that case the forward pass yields $W_A(a) = 0$ hence (15) holds.

For links a such that $\Omega_r^a = 1$:

o either their initial node is $B_a = r$ and the forward pass yields $W_A(a) = A(a) = \exp(-\theta T_a)$, hence (15) holds.

o or we can assume that (15) holds for all links a' such that $C_r^0(B_{a'}) < C_r^0(B_a)$. By recursion, since (15) holds for those links a' that precede a, (11) holds for $n = B_a$ and the forward pass computes $W_A(a)$ as

$W_A(a) = \exp(-\theta T_a)\Gamma_N(r, B_a) = \Gamma_A(r, E_a, a)$, owing to (13) applied to $n_1 = r$ and $n_2 = E_a$, since $\Gamma_A(B_a, E_a, a) = \exp(-\theta T_a)$ if $\Omega_r^a = 1$.

Lemma 3.

$$\sum_{k \in L_r(r,s)} \exp(-\theta T_{rs}^k) = W_N(s) \tag{17}$$

$$\frac{1}{\theta} \sum_k f_{rs}^k \log\left(\frac{f_{rs}^k}{q_{rs}}\right) = -\left\{ \sum_{k \in L_r(r,s)} f_{rs}^k T_{rs}^k \right\} - \frac{q_{rs}}{\theta} \log\left(\sum_{k \in L_r(r,s)} \exp(-\theta T_{rs}^k) \right) \tag{18}$$

Proof of Lemma 3.

Eqn. (17) is a mere application of (16), useful in computing the logit level-of-service (Williams 1977) $S_{rs}(\mathbf{T}) = -\frac{1}{\theta}\log(\sum_k \exp(-\theta T_{rs}^k))$: it holds that $S_{rs}(\mathbf{T}) = -\frac{1}{\theta}\log(W_N(s))$. In the variable demand case, S_{rs} must be evaluated so as to supply the demand function D_{rs} with its argument and compute the trip rate from r to s as $Q_r(s) = D_{rs}(S_{rs})$. Equation (18) stems

from (2) and from $\log(f_{rs}^k / q_{rs}) = -\theta T_{rs}^k - \log(\sum_{k' \in L_r(r,s)} \exp(-\theta T_{rs}^{k'}))$.

Combined with (17), (18) provides an easy way to compute the entropic part of the primal objective function of the STOCH3 model: summing over the O-D pairs, we get that:

$\frac{1}{\theta} \sum_{rsk} f_{rs}^k \log(f_{rs}^k / q_{rs}) = -\sum_{rsk} f_{rs}^k (T_{rs}^k - S_{rs}(\mathbf{T}))$

$= -\sum_a T_a F(a) + \sum_{rs} q_{rs} S_{rs}(\mathbf{T})$

since $\sum_a T_a F(a) = \sum_{rs} \sum_{k \in L_r(r,s)} f_{rs}^k \{\sum_a \delta_{rs}^{ak} T_a\} = \sum_{rs} \sum_{k \in L_r(r,s)} f_{rs}^k T_{rs}^k$.

Lemma 4.

$$X_A(a) = \sum_s \sum_{k \in L_r(r,s)} \delta_{rs}^{ak} f_{rs}^k \quad \text{where} \quad f_{rs}^k = q_{rs} \frac{\exp(-\theta T_{rs}^k)}{\sum_{k' \in L_r(r,s)} \exp(-\theta T_{rs}^{k'})} \tag{19}$$

Proof of Lemma 4.

As the variables X_A and X_N depend linearly on the variables $Q_r(s)$ (recall the definition of the backward pass), it is sufficient to show that (19) holds when there is only one destination node s. Let (19s) be the same condition with one destination s.

We will give a recursive proof of (19s) considering the links a in the order of decreasing costs $C_r^0(E_a)$.

(19s) holds for all links a such that $C_r^0(E_a) \geq C_r^0(s)$, since in that case there is no reasonable path from r to s that is comprised of a, and the backward pass yields $X_A(a) = 0$, as does (19s).

Secondly we check (19) for the links a with final node $E_a = s$: from (16) and (17),

$$X_A(a) = q_{rs} W_A(a) / W_N(s) = \sum_{k \in L_r(r,s)} \delta_{rs}^{ak} q_{rs} \frac{\exp(-\theta T_{rs}^k)}{\sum_{k' \in L_r(r,s)} \exp(-\theta T_{rs}^{k'})}.$$

Finally, assuming that (19s) holds $\forall a'$ such that $C_r^0(E_{a'}) > C_r^0(E_a)$, we demonstrate (19s) for link a. We notice that (19s) holds for every link a' with initial node $B_{a'} = E_a$, because either $C_r^0(E_{a'}) > C_r^0(B_{a'})$, or $C_r^0(E_{a'}) \le C_r^0(B_{a'})$ and $\Omega_r^{a'} = 0$, hence the backward pass yields $W_A(a') = 0 = X_A(a')$ as in (19s) since in that case no reasonable path can include a'. Thus, the backward pass yields:

$$X_A(a) = \left\{ \sum_{a';B_{a'}=E_a} \sum_{k \in L_r(r,s)} \delta_{rs}^{a'k} q_{rs} \frac{\exp(-\theta T_{rs}^k)}{\Gamma_N(r,s)} \right\} \frac{\Gamma_A(r,E_a,a)}{\Gamma_N(r,E_a)}$$

$$= \frac{q_{rs}}{\Gamma_N(r,s)} \sum_{a';B_{a'}=E_a} \Gamma_A(r,s,a') \frac{\Gamma_A(r,E_a,a)}{\Gamma_N(r,E_a)}.$$

From (13) we know that $\Gamma_A(r,s,a') = \Gamma_N(r,E_a)\Gamma_A(E_a,s,a')$,

and from its definition, $\Gamma_N(E_a,s) = \sum_{a';B_{a'}=E_a} \Gamma_A(E_a,s,a')$, hence:

$$X_A(a) = \frac{q_{rs}}{\Gamma_N(r,s)} \Gamma_N(E_a,s)\Gamma_A(r,E_a,s)$$

$$= \frac{q_{rs}}{\Gamma_N(r,s)} \Gamma_N(E_a,s)\Gamma_A(r,s,a) \quad \text{from (14).}$$

This last expression is identical to (19s), because $f_{rs}^k = q_{rs} \exp(-\theta T_{rs}^k) / \Gamma_N(r,s)$.

Summing over indices s the equation (19s) gives (19).

Fig. 1. Proportion of travellers that choose path 1

as a function of θ and the time difference T_2-T_1 (binary case).

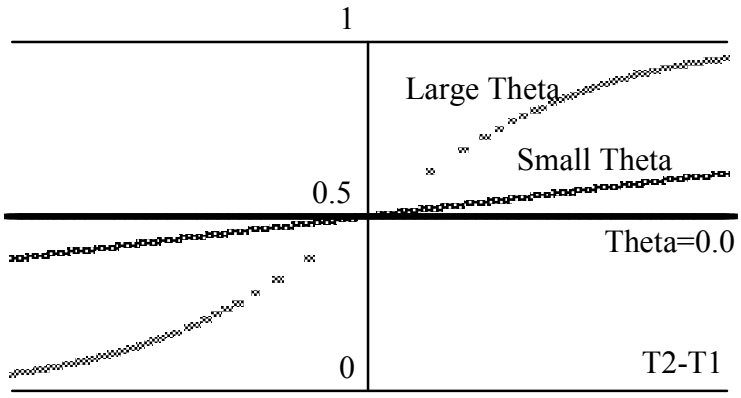

Fig. 2. Convergence rates of the three algorithms.