AN IMPLICIT PATH ENUMERATION MODEL AND ALGORITHM FOR DYNAMIC TRAFFIC ASSIGNMENT WITH CONGESTION SPILLBACK

Lorenzo Meschini, Giuseppe Bellei, Guido Gentile, Natale Papola
Dipartimento di Idraulica Trasporti e Strade, University of Rome "La Sapienza", Italy
Phone: +39 06 44585737 E-mail: lorenzo.meschini@uniroma1.it

Abstract

A new continuous formulation of Dynamic Traffic Assignment, where a user equilibrium is expressed as a fixed point problem in terms of arc flow temporal profiles, is proposed. The precise aim of the paper is to integrate spillback modelling into an existing formulation of DTA, based on implicit path enumeration, which is capable of representing explicitly the formation and dispersion of vehicle queues on the road network, but allows the queue length to overcome the arc length. Specifically, we take into account the interaction among network links upstream and downstream road intersections deriving from time varying entering and exiting arc capacities due to vehicle queue spillovers, which is equivalent to introducing constraints on the queue lengths. To achieve this extension we introduce a sequence of models, that is the arc entering capacity model, the arc exiting capacity model, and the arc travel time model for time varying exiting capacity, describing, for given turning flow temporal profiles, the dynamic of the network nodes and arcs, and capable of representing the propagation of congestion among contiguous road links. Some numerical examples are devised to appreciate the relevant effect of spillback modelling in the context of DTA.

Keywords: Dynamic traffic assignment; Queue spillback; Implicit path enumeration
Topic Area: D3 Integrated Supply / Demand Modelling

1 Introduction

During the last decade Dynamic Traffic Assignment has been one of the most active fields in transport modelling. Indeed, the classical static framework of traffic assignment is often recognized to be an improper tool for the analysis of highly congested road networks where the formation and the dispersion of vehicle queues plays a decisive role. This is particularly true for those applications (variable massage signs, on-line information to drivers, ramp metering, lane variable usage) where the estimation of travel times is the desired output and not only the flow pattern is relevant. At the same time, the continuous growth of computer performances induces to think that DTA models could be used in the near future also for planning purposes.

Due to the temporal dimension added, a new problem arises characterizing DTA, that is: loading the network with given demand flows on specified paths, in such a way that the resulting arc flow temporal profiles are consistent through an arc performance model with the corresponding travel time temporal profiles. This aspect, referred to as Continuous Dynamic Network Loading (CDNL), is so important that in the literature much attention has been dedicated to its specific analysis [see for instance Xu *et al.*, 1999]. In fact, this one is the only problem to be solved in those cases, such as highway networks and urban corridors, where the path choice is not relevant. The relative simplicity of these network structures has induced researchers to describe the temporal dimension introducing "short time intervals"

(1-10 sec), thus exploiting in their models the fact that a vehicle entering a given arc during a certain time interval will exit that arc not earlier than the next time interval.

Many existing DTA models are indeed an attempt to combine the CDNL and a path choice model, usually consisting of a dynamic shortest path procedure, into a dynamic equilibrium, thus permitting the analysis of more general networks. However, this approach, while perfectly correct from the modelling point of view, becomes often unpractical form the algorithmic point of view, due to the great amount of computer time and memory that is required.

In a recent paper [Bellei, Gentile and Papola, 2002] we have proposed a new continuous formulation of Dynamic Traffic Assignment (DTA), where a user equilibrium is expressed as a fixed point problem in terms of arc flow temporal profiles. There, it is shown that, by extending to the dynamic case the concept of Network Loading Map (NLM), is no more needed to introduce the CDNL as a sub-problem of DTA in order to ensure the temporal consistency of the supply model. On this basis it is possible to devise efficient assignment algorithms, whose complexity is equal to the one resulting in the static case multiplied by the number of time intervals in which the period of analysis is divided.

With specific reference to a Logit route choice model where only efficient paths are considered, an implicit path enumeration formulation of DTA is devised exploiting the concepts of arc conditional probabilities and node satisfactions, and a specific network loading procedure is also obtained as an extension of Dial's algorithm; then, the fixed point problem is solved through an accelerated averaging procedure, called Bather's Method. Another important feature of the proposed model lays in the fact that it does not exploit the acyclic graph characterizing the corresponding discrete version of the problem. Specifically, we do not introduce the hypothesis that the largest time interval must be shortest than the smallest free flow speed arc travel time. This way, at the algorithm level, it is possible to define "long time intervals" (5-10 min) and this allows to solve large instances of the problem with reasonable computing time and memory [Gentile and Meschini, 2003].

The fixed point problem is formalized by combining the NLM, yielding the arc inflow temporal profiles corresponding to given arc travel time and cost temporal profiles, and an arc performance function, yielding the arc travel time and cost temporal profiles corresponding to given inflow temporal profiles, capable of representing explicitly the formation and dispersion of vehicle queues, but allowing the queue length to overcome the arc length (vertical queues).

The aim of this paper is to extend this formulation in order to take into account the interaction among network links upstream and downstream road intersections deriving from time varying entering and exiting arc capacities due to vehicle queue spillovers, which is equivalent to introducing constraints on the queue lengths.

This approach is consistent with the definition, provided by [Adamo *et al.*, 1999], of the *spillback* phenomenon as a hypercritic flow state, either propagating backward from the final section of an arc and reaching its initial section, or originating on the latter, that reduces the exiting capacities of the arcs belonging to its backward star and eventually influences their flow states. Specifically, it shall be pointed out that this paper aims at representing a "polite" spillback, where users that cannot enter a given arc due to the presence of a downstream queue do not occupy the intersection, but wait on the upstream arc until the space necessary to their entrance becomes available on the downstream arc.

To achieve this extension we introduce a sequence of models describing, for given turning flow temporal profiles, the dynamic of the network nodes and arcs, and capable of representing the propagation of congestion among contiguous road links. Specifically, the Network Performance Model (NPM) will be articulated in a sequence of three models, namely the *arc entering capacity model*, the *arc exiting capacity model*, and the *arc travel time model for time varying exiting capacity*.

The paper is organized as follows. First, we present the formulation of DTA with spillback. Second, we describe the chain of sub-models making up the NPM. Third, we present the implicit path formulation of the NLM and the resulting dynamic equilibrium model. Fourth, we propose an heuristic solution algorithm. Fifth, we devise some numerical examples to appreciate the relevant effect of spillback modelling. Finally some short conclusions and consideration on future research perspectives are drawn.

2 Formulation of DTA with spillback

Limiting our attention to macroscopic flow modelling and thus neglecting the microscopic approaches to the problem, very few works in the literature achieve in representing spillback within the CDNL; among these [Daganzo, 1994 and 1995; Adamo *et al.*, 1999]. The above models satisfy the FIFO rule and can be applied to networks with many origins and destinations, although they require "short time intervals" which results in a demanding algorithm applicable to limited size networks. The Cell Transmission Model is consistent with the Simplified Theory of Kinematic Waves, but requires also a thick spatial discretization. The formulation proposed by Adamo *et al.* is founded on link-based models, but requires an "event-based" time discretization which is intrinsically not handy to implement and resembles micro simulation.

In these CDNL models the information on path choice is given in a local form by exogenous node splitting rates, possibly specified for each destination. Thus, their employment in the context of DTA requires the temporal profiles of the splitting rates to be calculated endogenously. Although, the latter is not a direct output of the path choice model [Papageorgiu, 1990], and this may be one of the reasons for the lack of models addressing DTA with spillback without tracking flows on each path, as in (Lo and Szeto, 2002).

Here we adopt a different approach which consists in solving the non-separability of the performances due to queue spillovers, the consistency of path choice with the path costs resulting from the network loading, and the CDNL jointly, by iterating, through an averaging algorithm, a sequence of models which will be consistent among each other only when the convergence is achieved, that is at the equilibrium. In doing this, we will introduce the representation of the spillback phenomenon simply by simulating its effects on the performance model, thus avoiding any interaction with the loading model. Specifically, we will exploiting the idea of time varying capacities as a function of the current arc flow temporal profiles; the arc travel times become then a function of the time varying arc exiting capacities; while the NLM simply loads the demand flows on the available paths based on the arc conditional probabilities resulting from the route choice model consistently with the current arc travel times.

This way we will accomplish to introduce the representation of queue spillovers into the model proposed in [Bellei, Gentile and Papola, 2002], where the DTA is formalized and solved as a fixed point problem in terms of the arc inflow temporal profiles. Although, in order to model correctly the spillback phenomenon, we consider here the turning flows (that is, the inflows disaggregated by maneuver) as current variable of the fixed point problem, as these have a role in how the available capacity at a node is split among the upstream arcs. Figure 1 depicts the scheme of the proposed formulation. In the following sections we will formalize all the sub-models and then formulate DTA with spillback as a simple fixed point problem.

Note that, in the alternative, DTA can be thought as three fixed point problems one inside the other. Specifically, we would have the inner fixed point consisting of the CDNL, then the mid fixed point problem combining the inner problem with the capacity model, and finally the outer fixed point problem combining the mid problem with the path choice model.

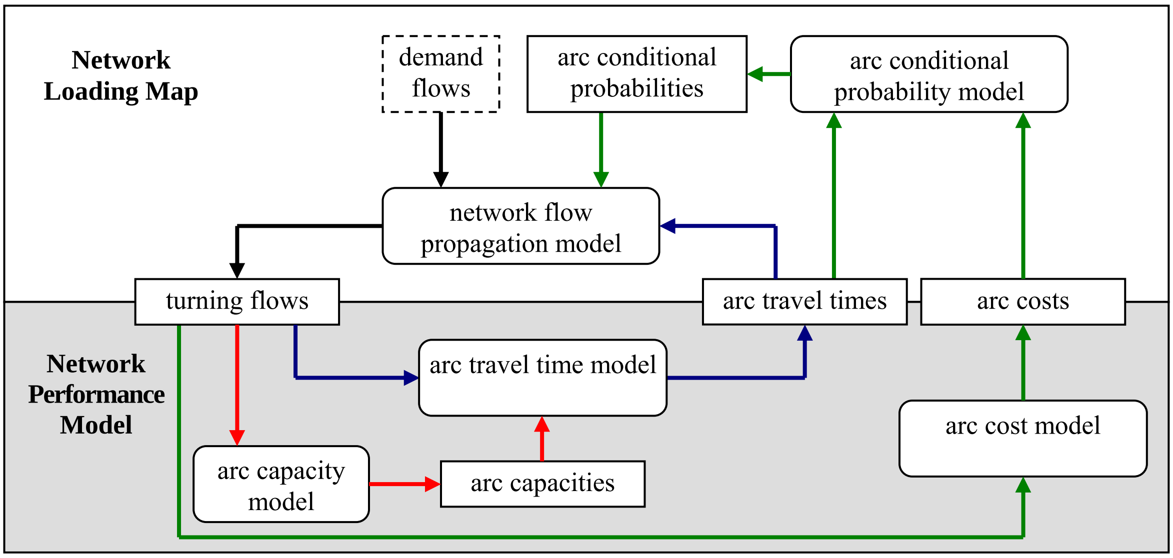

Figure 1. Scheme of the fixed point formulation of DTA with spillback.

3 Network performance model

In the following we will present the NPM referring to the generic elements (nodes and arcs) of the road network, while the origins and destinations of trips are regarded as special nodes, called centroids, that are assumed to be connected to the road network through special arcs, called connectors, characterized by infinite capacity.

The generic arc is modelled here in three parts: an *initial bottleneck*, a *final bottleneck* and a *running link*.

The time varying capacity of the initial bottleneck takes into account the physical capacity of the link and the effect of queues propagating backward on the arc itself that can reach the initial section and can thus induce spillback conditions. Specifically, it is set to a value that, given the current outflow, would limit the current inflow on the running link just in order to ensure that at any time the number of users on the link does not exceed its storage capacity and that the arc physical capacity is not overcome. Clearly, at the equilibrium the current inflows will not exceed the time varying arc entering capacities, because the implicit constraints on the queue lengths will be satisfied; however this is not true in general at a given step of the solving procedure.

The time varying capacity of the final bottleneck takes into account the hypercritic flow states, referred to as *queue*, that are generated either by a constant reduction of the physical capacity at the end of the arc due to the average effect of a road intersection, or by spillback conditions on the downstream arcs. The arc exiting capacities are obtained from the arc entering capacities based on flow conservation at nodes as a function of the turning lows and of the geometry of the intersection, synthetically expressed by the reduced arc physical capacities.

The running link models the congestion due to the interaction along the arc of vehicles travelling in hypocritical conditions; it consists of a homogeneous channel where flow states are determined on the basis of the Simplified Kinematic Wave theory assuming a certain *fundamental diagram*.

For simplicity, in the following we assume that the fundamental diagram of the generic arc a, with initial node $TL(a)$, final node $TL(a)$ and length L_a , has a triangular shape, as depicted in Figure 2, where Q_a is the arc physical capacity, V_a is the free flow speed, and w_a is the absolute value of the hypercritical wave speed. However, the extension to more complex flow models can be achieved based on the methods proposed in [Gentile, Meschini, Papola, 2003].

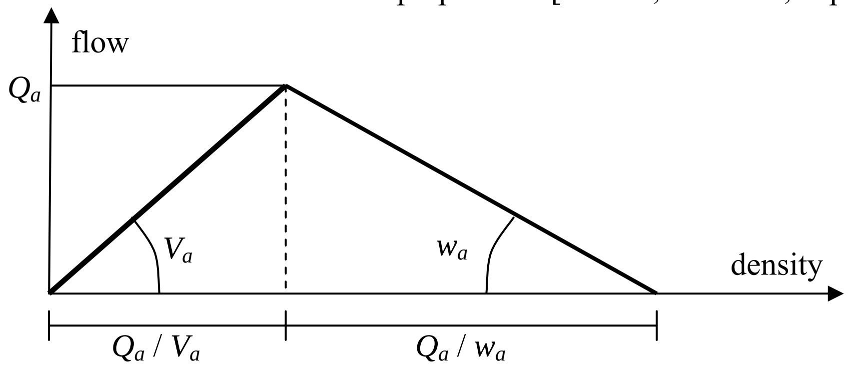

Figure 2. Fundamental diagram.

Within this framework, the NPM determines the temporal profiles of arc entering capacities, of arc exiting capacities, and finally of arc travel times, for given arc turning flow temporal profiles, as described in the Table 1 below.

Table 1. Sub-models input and output of the NPM.

model	**input**	**output**
arc entering capacities (separable)	arc inflow temporal profiles arc outflow temporal profiles arc physical capacities	arc entering capacity temporal profiles
arc exiting capacities (NOT separable)	arc physical capacities arc entering capacity temporal profiles arc turning flow temporal profiles	arc exiting capacity temporal profiles
arc performances (separable)	arc inflow temporal profiles arc exiting capacity temporal profiles	arc travel time temporal profiles

3.1 Arc entering capacity model

This model represents the effect on the arc entering capacity of queues generated on the arc final section due to a limited exiting capacity. Specifically, it determines the maximum inflow that maintains the queue length not longer than the arc length, and by comparison with the actual inflow it identifies the time intervals when the spillback phenomenon occurs.

Clearly, if the queue was rigid (only one hypercritical density) any hypercritical flow state occurring at the arc final section would propagate instantaneously to the arc initial section and thus for any instant when the queue exceeds the length of the arc, we would have that the entering capacity is equal to the outflow. However, in reality hypercritical flow states may occur at different densities (stop and go) and the propagation speed of the flow states is finite. This can be taken into account avoiding the explicit evaluation of the queue temporal profile, which is cumbersome since the speed and density of vehicles in the queue vary accordingly with the outflow. Instead, exploiting the analytical solution of the Simplified Kinematic Wave Theory based on cumulative flows proposed by [Newell, 1993], the model identifies the time instants when the outflow states assumed hypercritical (the potential queue), propagating backward along the arc, would reach the initial section.

Let $G_a(\tau)$ be the backward propagation of the cumulative outflow temporal profile $E_a(\tau)$ assumed hypercritical, representing the maximum cumulative flow temporal profile that could have entered the arc without violating the constraint on the queue length consistently with the current outflow pattern. We have in general [Gentile, Meschini, Papola, 2003]:

$$H_a(\tau) = E_a(\tau) - L_a \cdot \left(\frac{e_a(\tau)}{\mathrm{w}_a\left(e_a(\tau)\right)} - \mathrm{k}_a\left(e_a(\tau)\right) \right),$$

$$u_a(\tau) = \tau + \frac{-L_a}{\mathrm{w}_a\left(e_a(\tau)\right)}, \tag{1}$$

$$G_a(\tau) = \inf\left\{ H_a(\sigma) : u_a(\sigma) = \tau \right\},$$

where $u_a(\tau)$ is the instant when the backward kinematic wave generated by the outflow $e_a(\tau)$ at time τ considered hypercritical would reach the initial section, $H_a(\tau)$ is the cumulative flow that would be observed at time $u_a(\tau)$ in the initial section, while $\mathrm{w}_a(\cdot)$ and $\mathrm{k}_a(\cdot)$ express respectively the speed of the kinematic wave (which is negative for hypercritical flow states) and the density as a function of the flow depending on the fundamental diagram adopted. The latter formula in the above system of equations expresses the Newell-Luke minimum principle [Daganzo, 1997; Newell, 1993], stating that, when more than one kinematic wave reachs a point at a same time, the flow state yielding the minimum cumulative flow dominates the others. The solution of system (1) for the proposed triangular fundamental diagram is straightforward:

$$G_a(\tau) = E_a(\tau - \frac{L_a}{w_a}) + L_a \cdot Q_a \left(\frac{1}{w_a} + \frac{1}{V_a} \right). \tag{2}$$

On this basis, the entering capacity $\mu_a(\tau)$ can be set to $dG_a(\tau)/d\tau$, if the spillback constraint is active (that is, when $G_a(\tau)$ is not higher than $F_a(\tau)$), and to the physical capacity Q_a, otherwise:

$$\mu_a(\tau) = \begin{cases} \dfrac{dG_a(\tau)}{d\tau}, & \text{if } G_a(\tau) \le F_a(\tau); \\ Q_a, & \text{otherwise.} \end{cases} \tag{3}$$

Differentiating equation (2) we obtain:

$$\frac{dG_a(\tau)}{d\tau} = e_a(\tau - \frac{L_a}{w_a});$$

then, for the proposed triangular fundamental diagram equation (3) becomes:

$$\mu_a(\tau) = \begin{cases} e_a(\tau - \dfrac{L_a}{w_a}), & \text{if } G_a(\tau) \le F_a(\tau); \\ Q_a, & \text{otherwise.} \end{cases} \tag{4}$$

Note that equation (4) is consistent with the limitation to the inflow due to the arc physical capacities, because the outflow produced by the NLM is always not greater than the reduced physical capacity.

Figure 3 below shows graphically how the model works in this case: profile $G_a(\tau)$ is obtained by a rigid translation of profile $E_a(\tau)$ for $-L_a/w_a$ in time and for $L_a \cdot Q_a \cdot (1/w_a + 1/V_a)$ in value. When $G_a(\tau)$ is above $F_a(\tau)$, either the queue is shorter that the arc or it is zero and $\mu_a(\tau)$ is set to Q_a ; otherwise spillback occurs and $\mu_a(\tau)$ is set to $e_a(\tau - L_a/w_a)$.

Since it obviously holds:

$$f_a(\tau) = \sum_{b \in BS(TL(a))} \varphi_{ba}(\tau), \; F_a(\tau) = \int_0^\tau f_a(t) \cdot \mathrm{d}t,$$

$$e_a(\tau) = \sum_{b \in FS(HD(a))} \varphi_{ab}(\tau)\,,\; E_a(\tau) = \int_0^\tau e_a(t) \cdot \mathrm{d}t\,,$$

where $f_a(\tau)$ is the inflow at time τ and $\varphi_{ab}(\tau)$ is the turning flow from arc a to arc b at time τ, with $HD(a) = TL(b)$, we can express the arc entering capacity model in a compact form as follows:

$$\mu = \mu(\varphi\,;\mathbf{Q}). \tag{5}$$

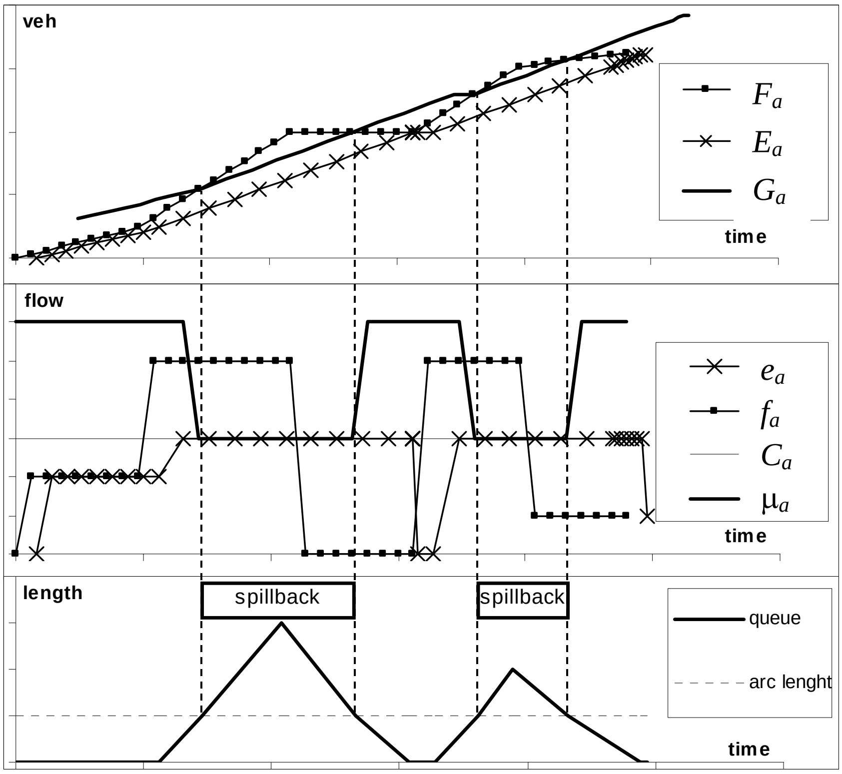

Figure 3. Determination of the entering capacity temporal profile for given arc flows.

3.2 Arc exiting capacity model

This model determines with reference to a given node, the time varying exiting capacities of the upstream arcs on the basis of the entering capacities of the downstream arcs and of the local turning flows. The resulting model is spatially non separable, because the exiting capacities of all the arcs belonging to the backward star of a same node are determined jointly; it is indeed the model that propagates the spillback congestion through the network.

In order to simplify the presentation we first consider only two typology of nodes: *merging* and *diversions*. In this case the turning flows are expressed directly through the arc inflows and outflows.

When considering a *merging*, the problem is how to split the available capacity $\mu_a(\tau)$ of an arc a among the outflows $e_b(\tau)$ of the arcs $b \in BS(TL(a))$ of its backward star. In principle, we assume that the available capacity is split proportionally to their reduced physical capacities $C_b \leq Q_b$ (although the capacity of connectors is infinite, their reduced capacity is finite). However, some of these arcs may not require all the capacity this way assigned to them as their outflow is lower than it. Then we can split the residual capacity among those arcs that would utilize more capacity if available (the arcs of the set $B \subseteq BS(TL(a))$ defined below). On this basis, the arc exiting capacities $\xi_b(\tau)$ are given by:

$$\xi_b(\tau) = \begin{cases} \dfrac{C_b}{\sum_{c\in B} C_c} \cdot \left[\mu_a(\tau) - \sum_{c\in BS(TL(a))\setminus B} e_c(\tau) \right], & \text{if } c \in B; \\ C_b, & \text{otherwise,} \end{cases} \tag{6}$$

$$B = \{ b \in BS(TL(a)) : e_b(\tau) \geq \xi_b(\tau) \} . \tag{7}$$

The separation set B satisfying jointly (6) and (7) can be obtained, starting from $BS(TL(a))$, by subtracting iteratively from the current set the arcs for which the entering capacity determined through (6) is higher than the outflow (at the most $|BS(TL(a))|$ steps are required). Clearly, if $\sum_b e_b(\tau) < \mu_a(\tau)$, then no spillback phenomenon is active and $B = \varnothing$.

When considering a *diversion*, the problem is to determine the most severe reduction of outflow from an arc a due to the entering capacities of the arcs $b\in FS(HD(a))$ of its forward star. Specifically, if the spillback phenomenon is active form one of these arcs, that is $f_b(\tau) \geq \mu_b(\tau)$, in order to satisfy its entering capacity, based on the FIFO rule we have to limit its inflow in such a way that the ratio between its entering capacity $\mu_b(\tau)$ and the exiting capacity $\xi_a(\tau)$ of arc a is equal to the ratio between the turning flow on arc b and the total outflow from arc a:

$$\frac{\mu_b(\tau)}{\xi_a(\tau)} = \frac{f_b(\tau)}{e_a(\tau)} .$$

Clearly if the spillback is active on more that one arc $b\in FS(HD(a))$, the exiting capacity $\xi_a(\tau)$ is set consistently with the most penalizing one. If no spillback is active, then the exiting capacity equals the reduced physical capacity. In general we have:

$$\xi_a(\tau) = \min \left\{ C_a \; , \; \mu_b(\tau) \cdot \frac{e_a(\tau)}{f_b(\tau)} \; \forall b \in FS(HD(a)) : f_b(\tau) \geq \mu_b(\tau) \right\} .$$

When considering a generic node x with both merging and diversion, the turning flows are to be considered explicitly as follows:

$$\xi_{ab}(\tau) = \begin{cases} \dfrac{C_a}{\sum_{c\in B(b)} C_c} \cdot \left[\mu_b(\tau) - \sum_{c\in BS(x)\setminus B(b)} \varphi_{cb}(\tau) \right], & \text{if } c \in B(b); \\ C_b, & \text{otherwise,} \end{cases} \quad a \in BS(x), b \in FS(x) ,$$

$$B(b) = \{ a \in BS(x) : \varphi_{ab}(\tau) \geq \xi_{ab}(\tau) \}, \quad b \in FS(x) ,$$

$$\xi_a(\tau) = \min \left\{ C_a \; , \; \xi_{ab}(\tau) \cdot \frac{e_a(\tau)}{\varphi_{ab}(\tau)} \; \forall b \in FS(x) : \varphi_{ab}(\tau) > \mu_b(\tau) \right\}, \quad a \in BS(x) ,$$

where $\xi_{ab}(\tau)$ denotes the maximum turning flow from arc $a\in BS(x)$ to arc $b\in FS(x)$.

The arc exiting capacity model can thus be expressed in the following compact form:

$$\xi = \xi(\varphi \, , \, \mu \, ; \, \mathbf{C}). \tag{8}$$

3.3 Arc travel time model for time varying exiting capacity

Once the temporal profile of the exiting capacity of each arc has been determined, it is possible to utilize the following link based macroscopic arc performance model to determine the arc travel time temporal profiles.

When there is no capacity reduction at the end of the arc, the running time $r_a(\tau)$ is in general a function of the arc inflow temporal profile [Gentile, Meschini, Papola, 2003]. However, in the particular case of a triangular fundamental diagram the running time is constant in time and equal to L_a / V_a .

We are interested here in determining, with reference to a bottleneck with time varying capacity $\xi_a(\tau)$ and constant running time $r_a(\tau)$, the exit time temporal profile $t_a(\tau)$ resulting from a given inflow $f_a(\tau)$ due to a potential "vertical" queue. The cumulative outflow temporal profile is given by:

$$E_a(\tau) = \min\left\{ F_a(\sigma - r_a(\tau)) + \int_\sigma^\tau \xi_a(t)\,\mathrm{d}t : \sigma \le \tau \right\}. \tag{9}$$

The above expression can be explained as follows. When there is no queue, the travel time is equal to the running time, so that the outflow at a given time τ is equal to the inflow at time τ-$r_a(\tau)$. If a queue begins at time σ, the outflow will follow from that time until the queue vanishes the shape of the exiting capacity temporal profile. Then, based on the Newell-Luke principle, the actual outflow at time τ is the one yielding the minimum value among all the possible outflow that would occur if the queue would begins at any previous instant $\sigma \le \tau$.

Based on the FIFO rule, the travel time temporal profile can be determined solving the following implicit equation:

$$E_a(t_a(\tau)) = F_a(\tau). \tag{10}$$

Note that, based on the monotonicity of the three profiles involved, the numerical computation of $t_a(\tau)$ starting from $E_a(\tau)$ and $F_a(\tau)$ is trivial.

Figure 4 show the graphical interpretation of the cumulative outflow temporal profile as the lower envelop of a family of curves related to the exiting capacities and of the cumulative inflow shifted forward in time by the running time, and the calculation of the exit time based on the cumulative inflow and outflow temporal profile, in the case of a triangular fundamental diagram.

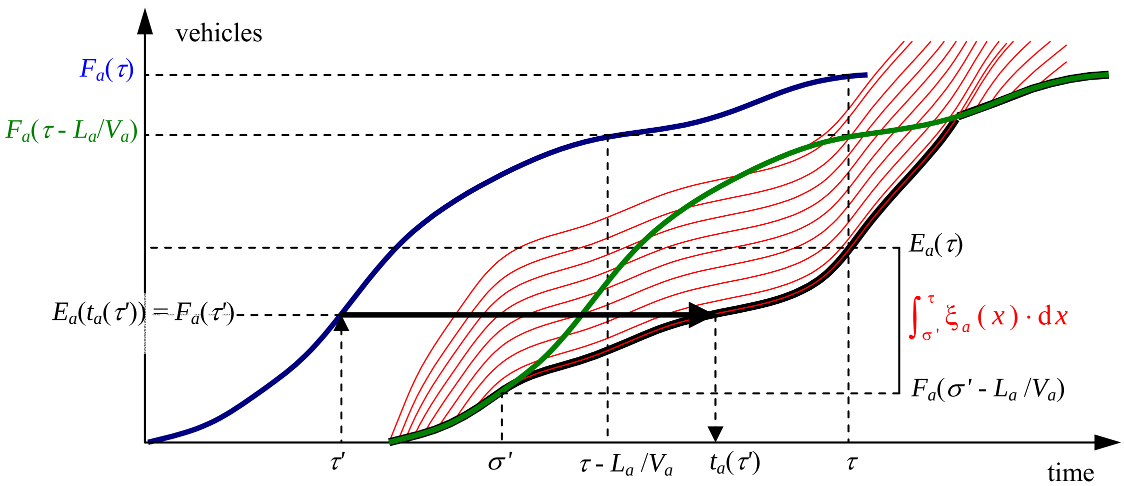

Figure 4. Bottleneck with time varying capacity.

Note that the proposed arc travel time model applies also to connectors; the exiting capacity of a connector entering a centroid is infinite, while that of a connector exiting a centroid is obtained using the arc exiting capacity model.

The arc travel time model for time varying exiting capacity can thus be expressed in the following compact form:

$$\mathbf{t} = \mathrm{t}(\varphi , \xi). \tag{11}$$

3.4 Arc cost model

The arc cost for users entering at time τ is simply assumed to be:

$c_a(\tau) = \eta\cdot(t_a(\tau) - \tau) + m_a(\tau)$,

where m_a is the temporal profile of the monetary cost, while η is the Value of Time.

In a compact form we have:

$\mathbf{c} = \mathrm{c}(\mathbf{t})$. (12)

4 Network loading map and dynamic equilibrium model

4.1 Implicit path formulation of the route choice and network loading models

In the following we briefly recall the implicit path formulation addressing both route choice and network flow propagation presented in [Bellei, Gentile, Papola, 2002], referring to users travelling to destination d. This is founded on the concepts of *arc conditional probability* and *node satisfaction*; the corresponding notation and definitions are introduced below:

$p_a^d(\tau)$ probability of using arc a, conditional on being at node $TL(a)$ at time τ;

$w_x^d(\tau)$ expected value of the maximum perceived utility at time τ relative to the paths K_{xd}, connecting node x to d.

In the cited paper it is proved that the following expression of the node satisfaction and of the arc conditional probabilities are consistent with a Logit route choice model where only efficient paths are considered:

$$w_x^d(\tau) = \theta\cdot\ln\left(\sum_{a\in FSE(x,d)} \exp\left(\frac{-c_a(\tau) + w_{HD(a)}^d(t_a(\tau))}{\theta}\right)\right), \tag{13}$$

$$p_a^d(\tau) = \exp\left(\frac{-c_a(\tau) + w_{HD(a)}^d(t_a(\tau)) - w_{TL(a)}^d(\tau)}{\theta}\right), \tag{14}$$

where $FSE(x, d)$ is the efficient forward star of node x toward destination d, $w_d^d(\tau) = 0$, and $p_a^d(\tau) = 0$ if $a\notin FSE(HD(a), d)$. The solution of the system of equations (13) and (14), which constitutes the route choice model, can be expressed in compact form as:

$\mathbf{w} = \mathrm{w}(\mathbf{c}, \mathbf{t})$, (15)

$\mathbf{p} = \mathrm{p}(\mathbf{w}, \mathbf{c}, \mathbf{t})$. (16)

The turning flow from arc $a\in BS(x)$ to arc $b\in FS(x)$ directed to d is given by the outflow $e_a^d(\tau)$ of arc a directed to d multiplied by the arc conditional probability $p_b^d(\tau)$; then, summing up the contribution of each destination, we have:

$\varphi_{ab}(\tau) = \sum_{d\in C} p_b^d(\tau)\cdot e_a^d(\tau)$. (17)

The inflow $f_a^d(\tau)$ of arc a directed to d is given by the arc conditional probability $p_a^d(\tau)$ multiplied by the flow entering node $TL(a)$; the latter is given, in turn, by the sum of the outflows of the arcs belonging to the efficient backward star $BSE(TL(a), d)$ of $TL(a)$ toward destination d, and of the demand flow from $TL(a)$ to d, which is null when $TL(a)$ does not belong to the set of centroids C, that is:

$f_a^d(\tau) = p_a^d(\tau)\cdot[d_{TL(a)d}(\tau) + \sum_{b\in BSE(TL(a), d)} e_b^d(\tau)]$. (18)

Given the travel time temporal profile of arc a, the outflow is related to the inflow temporal profile as follows:

$e_a^d(t_a(\tau))\cdot \mathrm{d}t_a(\tau)/\mathrm{d}\tau = f_a^d(\tau)$, (19)

where the weight $\mathrm{d}t_a(\tau)/\mathrm{d}\tau$ stems from the fact that arc travel times vary over time, so that those users who enter the arc at a certain rate, in general, exit the arc at a different rate, which is higher, if the arc travel time is decreasing, and lower, otherwise (for details, see for example Cascetta, 2001).

The solution of the system of equations (17), (18) and (19), which constitutes the network flow propagation model, is formally expressed in compact form as:

$\varphi = \omega(\mathbf{p}, \mathbf{t}\ ; \mathbf{d})$. (20)

4.2 Equilibrium model with spillback

In the previous sections we have developed a sequence of mathematical models reproducing the within-day network dynamic of road traffic including the spillback phenomenon. These models naturally lead to a fixed point formulation of DTA based on the system between the NLM and the NPM, as depicted in Figure 5 below.

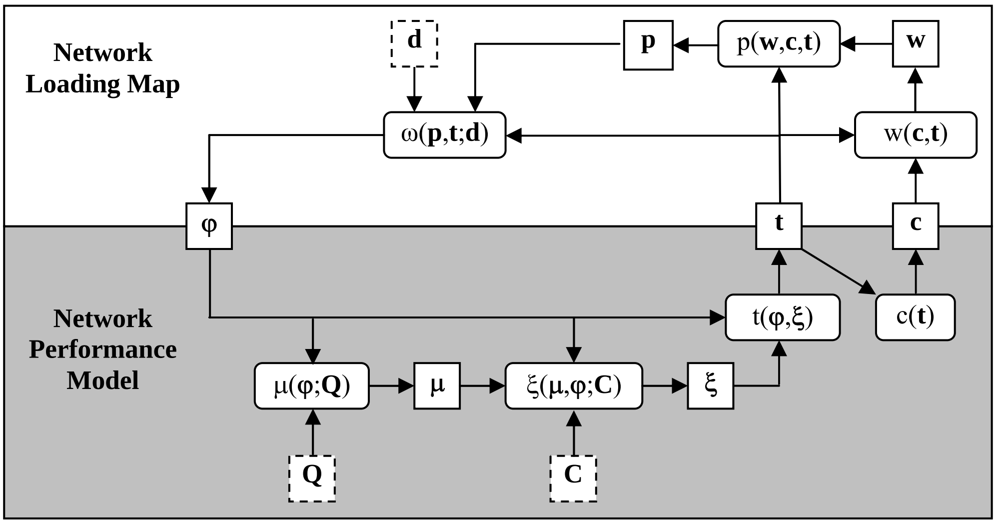

Figure 5. Variables and models of the fixed point formulation of DTA with spillback.

In analogy with the static case, the NLM is a functional relation which allows determining, for given arc performances, an arc flow pattern consistent with the demand flows through the path choice model. Combining (15) and (16) with (20) we obtain the following implicit path formulation of the NLM:

$\varphi = \omega(\mathrm{p}(\mathrm{w}(\mathbf{c}, \mathbf{t}), \mathbf{c}, \mathbf{t}), \mathbf{t}\ ; \mathbf{d})$. (21)

The NPM is the functional relation between the turning flows and the travel times, which can be obtained by combining (5), (8) and (11) as follows:

$\mathbf{t} = \mathrm{t}(\varphi\ , \xi(\varphi\ , \mu(\varphi\ ; \mathbf{Q})\ ; \mathbf{C}))$. (22)

On this basis, DTA can be formalized as a fixed-point problem in terms of turning flow temporal profiles by substituting into the NLM (21), the NPM equations (22) and (12).

5 Solution algorithm

In this section, we present only the algorithm solving the NPM, referring to [Bellei, Gentile and Papola, 2002] for the algorithm solving the NLM and the fixed point problem resulting from their combination.

In order to implement the proposed mathematical model, the period of analysis is divided into n time intervals identified by the sequence of instants $(\tau^{0}, \ldots, \tau^{i}, \ldots, \tau^{n})$. By convention it is $\tau^{-1} = -\infty$ and $\tau^{n+1} = \infty$.

In the following we assume to approximate the generic temporal profile x_j through either a piece-wise constant or a piece-wise linear function defined by the values taken at such instants, so that for the two cases, in that order, we have:

$x_j(\tau^{0}) = x_j^{0}\ ,\ x_j(\tau) = x_j^{i},\ \tau\in(\tau^{i-1}, \tau^{i}],\ i = 1, \ldots, n,$ (23)

$x_j(\tau^0) = x_j^0$, $x_j(\tau) = x_j^{i-1} + (\tau - \tau^{i-1}) \cdot (x_j^i - x_j^{i-1}) / (\tau^i - \tau^{i-1})$, $\tau \in (\tau^{i-1}, \tau^i]$, $i = 1, \ldots, n$. (24)

Specifically, the temporal profiles of the flows are assumed piece-wise constant, while those of the other variables are assumed piece-wise linear.

The state of the network at time τ^0 is assumed to be known; here, without loss of generality, an initially unloaded network is considered.

The following procedure identifies a method for computing the arc entering capacity based on the model described in section 3.1 for the generic arc $a \in AS$ which is not a connector.

for each $a \in AS$
    $F_a^0 = 0$
    $E_a^0 = 0$
    $H_a^0 = L_a \cdot Q_a \cdot (1/w_a + 1/V_a)$
    for $i = 1, \ldots, n$
        $F_a^i = F_a^{i-1} + f_a^i \cdot (\tau^i - \tau^{i-1})$
        $E_a^i = E_a^{i-1} + e_a^i \cdot (\tau^i - \tau^{i-1})$
        $H_a^i = E_a^i + L_a \cdot Q_a \cdot (1/w_a + 1/V_a)$
    next i
next a
for each $a \in AS$
    $G_a^0 = 0$
    $H_a^{-1} = 0$
    $j = 0$
    for $i = 1, \ldots, n$
        do until $\tau^{j-1} + L_a / w_a < \tau^i \leq \tau^j + L_a / w_a$
            $j = j+1$
        loop
        $G_a^i = H_a^{j-1} + (H_a^j - H_a^{j-1}) \cdot (\tau^i - \tau^{j-1} - L_a / w_a) / (\tau^j - \tau^{j-1})$
        if $G_a^i > F_a^i$ and $G_a^{i-1} > F_a^{i-1}$ **then**
            $\mu_a^i = Q_a$
        else
            $\mu_a^i = (G_a^i - \min\{G_a^{i-1}, F_a^{i-1}\}) / (\tau^i - \tau^{i-1})$
        end if
    next i
next a

The following procedure identifies a method for computing the arc exiting capacity based on the model described in section 3.2 for the generic node x which is not a centroid.

for $i = 1, \ldots, n$
    for each $x \in N \setminus C$
        for each $b \in FS(x)$
            $B = BS(x)$
            return 1
            $CB = 0$, $FB = 0$
            for each $c \in BS(x)$
                if $c \in B$ **then**
                    $CB = CB + C_c$
                else

$FB = FB + \varphi_{cb}{}^{i}$
end if
next c
for each $a \in BS(x)$
if $a \in B(b)$ **then**
$\xi_{ab}{}^{i} = C_a / CB \cdot (\mu_b{}^{i} - FB)$
else
$\xi_{ab}{}^{i} = C_b$
end if
if $\varphi_{ab}{}^{i} < \xi_{ab}{}^{i}$ **then**
$B(b) = B(b) \setminus a$
goto 1
end if
next a
next b
for each $a \in BS(x)$
$\xi_a{}^{i} = C_a$
for each $b \in FS(x)$
if $\xi_{ab}{}^{i} \leq \varphi_{ab}{}^{i}$ and $\xi_a{}^{i} > \xi_{ab}{}^{i} \cdot e_a{}^{i} / \varphi_{ab}{}^{i}$ **then**
$\xi_a{}^{i} = \xi_{ab}{}^{i} \cdot e_a{}^{i} / \varphi_{ab}{}^{i}$
end if
next b
next a
next x
next i

The following procedure identifies a method for computing the exit time based on the model described in section 3.3 for the generic arc $a \in AS$ which is not a connector. Figure 6 depict the method graphically.

for each $a \in AS$
$E_a{}^{0} = 0$
$j = 1$
$F_a{}^{n+1} = F_a{}^{n}$
for $i = 1, \ldots, n$
do until $\tau^{j-1} - L_a / V_a < \tau^{i} \leq \tau^{j} - L_a / V_a$
$j = j+1$
loop
$E_a{}^{i} = F_a{}^{j-1} + (F_a{}^{j} - F_a{}^{j-1}) \cdot (\tau^{i} - \tau^{j-1} - L_a / V_a) / (\tau^{j} - \tau^{j-1})$
next i
for $i = 1, \ldots, n$
$TE = E_a{}^{i-1} + \xi_a{}^{i} \cdot (\tau^{i} - \tau^{i-1})$
if $E_a{}^{i} > TE$ **then**
$E_a{}^{i} = TE$
loop
next i
$E_a{}^{n+1} = E_a{}^{n}$
$j = 1$
$t_a{}^{0} = \tau^{0} + L_a / V_a$

for $i = 1, \ldots, n$
  do until $E_a^{j-1} < F_a^i \leq E_a^j$
    $j = j+1$
  loop
  if $E_a^j > E_a^{j-1}$ **then**
    $t_a^i = \tau^{j-1} + (F_a^i - E_a^{j-1}) \cdot (\tau^j - \tau^{j-1}) / (E_a^j - E_a^{j-1})$
    if $t_a^i < \tau^i + L_a / V_a$ **then**
      $t_a^i = \tau^i + L_a / V_a$
    end if
  else
    $t_a^i = \tau^i + L_a / V_a$
  end if
next i
next a

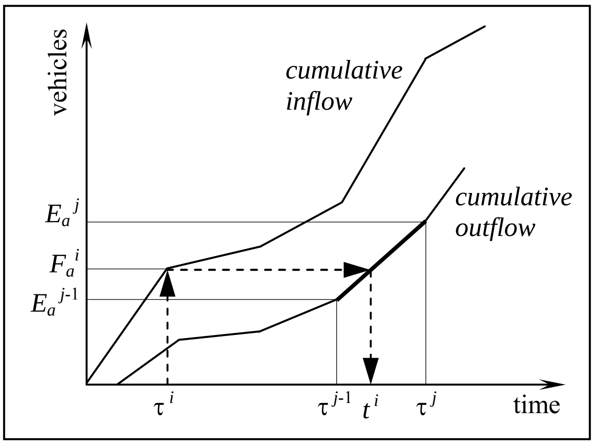

Figure 6. Exit time for given piece-wise linear cumulative inflow and outflow.

6 Numerical examples

In order to investigate the behavior of the proposed model, we will analyze three simple examples, which present intuitive solutions and could be solved also in closed form. In these examples we consider 100 intervals of 60 sec, assuming $V = 20$ m/s and $w = 0.25 \cdot V$.

The first two examples refer to the simple intersection depicted in Figure 7 below.

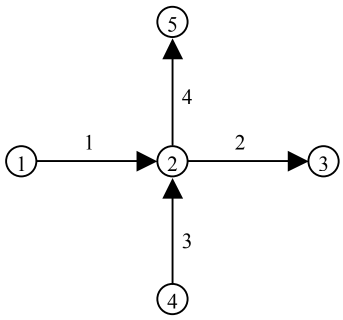

Figure 7. The network of example 1 and 2.

Assume first that the reduced capacity are: $C_1 = C_3 = 2000$ veh/h, $C_2 = C_4 = 900$ veh/h; and that there is a constant demand during the first 33 minutes of simulation: $d_{13} = 1200$ veh/h, $d_{43} = 400$ veh/h and $d_{45} = 600$ veh/h. The output of the two models with and without spillback are presented in Figure 8 below, showing that the results are quite different both in terms of flows and travel times. Without spillback, the congestion is concentrated on arc 2 and does not spread upstream: the travel time for the other arcs remains at the uncongested level, while the travel time of arc 2 grows until the end of the demand, since its inflow is not bounded. As a result, the travel time from node 4 to node 3 (arcs 3 and 2) gets very high; instead, the demand from node 4 to node 5 (arcs 3 and 4) travels at the free flow speed and is not involved in the congestion phenomenon. With spillback, the travel time on arc 2 has an upper bound, since congestion is transferred upstream as soon as the queue reaches the initial section of the arc; from that moment, the inflow on arc 2 drops to the outflow level, that is 900 veh/h. Since the outflow from arc 3, which is at the most equal to d_{43} , is lower than $C_2 \cdot C_4/(C_1+C_4)$ = 450 veh/h, arc 3 is not conditioned by the spillback phenomenon, and the congestion propagates only on arc 1. As a result, the travel time from node 4 to node 3 is much lower than the one calculated without spillback, since most of the queue develops now on arc 1 and only in part on arc 2.

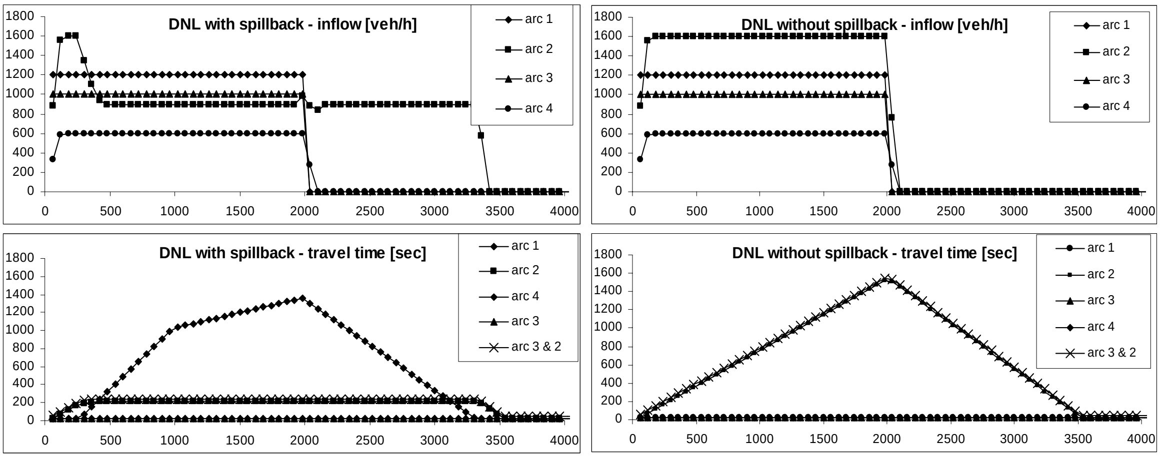

Figure 8. Output of the first example.

Assume now that the demand d_{43} is increased up to 600 veh/h. The output of the two models with and without spillback are presented in Figure 9 below. This example is very similar to the previous one, except for the fact, that now d_{43} is higher than $C_2 \cdot C_4/(C_1+C_4)$. Then, the congestion on arc 2 spills back on both arcs 1 and 3, and thus their travel time increases substantially. As a result, also the flow traveling from node 4 to node 5 is delayed, despite it does not use arc 2. Clearly, the model without spillback is unable to represent this effect.

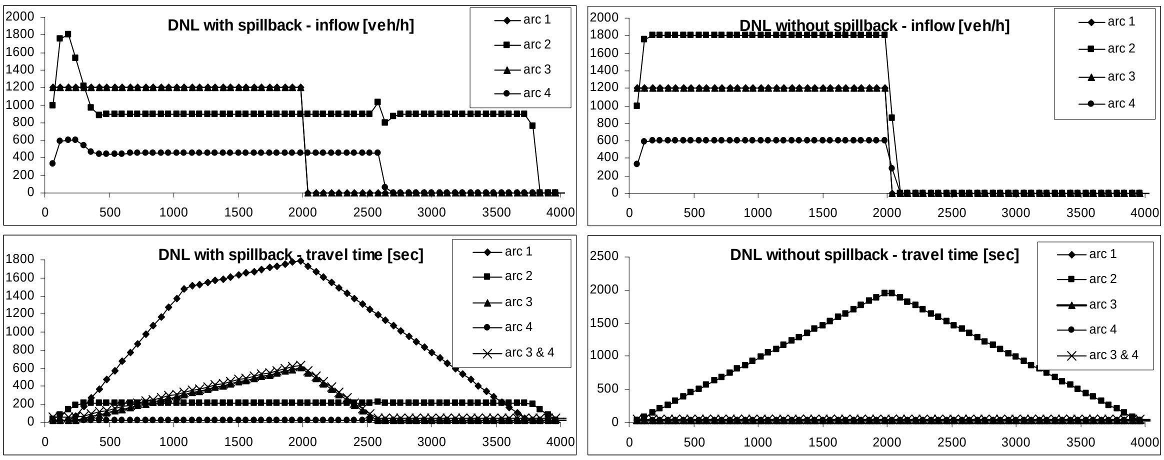

Figure 9. Output of the second example.

In the third example we consider the Braess network depicted in Figure 10 below with the aim of showing the effect of spillback on path choice. We assume the following reduced capacities: C_1 = C_5 =2000 veh/h, C_2 = C_3 = C_4 =1000 veh/h; and that there is a constant demand during the first 33 minutes of simulation: d_{14} = 2300 veh/h.

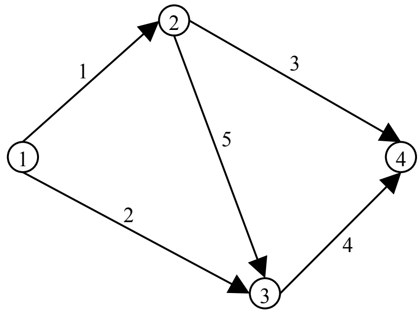

Figure 10. The network of example 3.

The output of the two models with and without spillback are presented in Figure 11 below. Without spillback, the congestion is located only on arcs 3 and 4, so that on all the paths between node 1 and node 4 the queue is equal; path 1-5-4 has few users, since it appears to be not convenient with respect to paths 2-4 and 1-3. If the spillback congestion is represented, however, the queue propagates toward arcs 1 and 2 from arcs 3 and 4, where, after an initial growth, it remains constant and equal to the arc length; path 1-5-4 becomes now competitive, as it implies a longer route but a shorter queue.

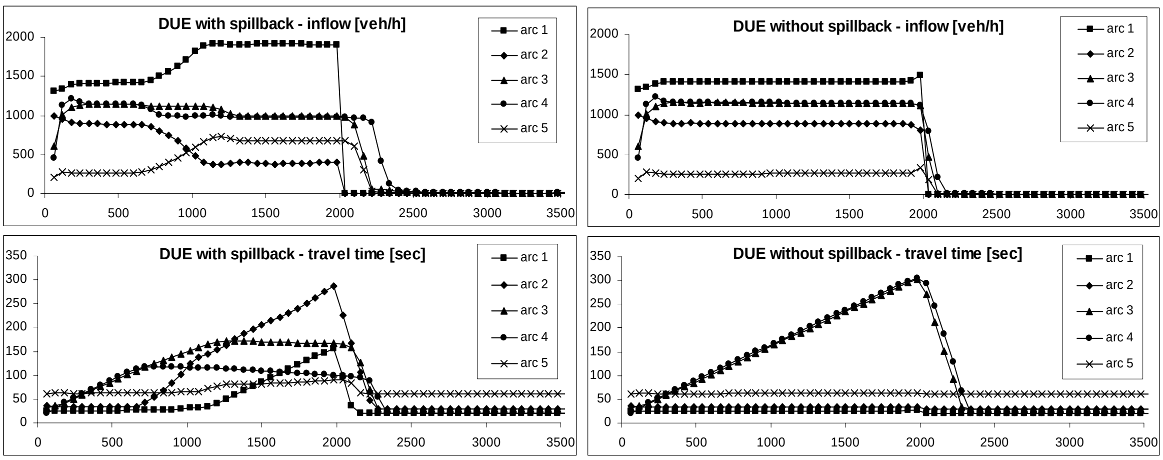

Figure 11. Output of the third example.

7 Conclusions

The proposed model has been applied to several elementary networks in order to verify the correspondence with the results obtained by other authors [Daganzo, 1998]. Particularly, interesting is the comparison of the DTA model with spillback and the previous model without it. More work is ongoing to improve the convergence of the solution algorithm and to test the method on a real network.

References

Adamo V., Astarita V., Florian M., Mahut M., Wu J.H., 1999. Modelling the spill-back of congestion in link based dynamic network loading models: a simulation model with application. in Proceedings of the 14th International Symposium on Transportation and Traffic Theory, ed. A. Ceder, Elsevier Science, Amsterdam, The Netherlands.

Bellei G., Gentile G., Papola N., 2002. A within-day dynamic traffic assignment model for urban road networks. accepted for publication on Transportation Research B.

Cascetta E., 2001. Transportation systems engineering: theory and methods, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Daganzo C.F., 1994. The cell transmission model: a dynamic representation of highway traffic consistent with hydrodynamic theory. Transportation Research B 28, 269-287.

Daganzo C.F., 1995. The cell transmission model, part II: network traffic. Transportation Research B 29, 79-93.

Daganzo C.F., 1997. Fundamentals of transportation and traffic operations, Pergamon, Oxford, UK, chapter 4.

Daganzo C.F., 1998. Queue spillovers in transportation networks with route choice. Transportation Science 32, 3-11.

Gentile G., Meschini L., 2003. A time-continuous implicit path formulation and algorithm for the Dynamic User Equilibrium problem. presented at AIRO Winter 2003, Champoluc, Italy.

Gentile G., Meschini L., Papola N., 2003. Macroscopic arc performance models with capacity constraints for within-day dynamic traffic assignment. submitted for publication to Transportation Research B.

Newell G.F., 1993. A simplified theory of kinematic waves in highway traffic, part I: general theory; part II: queuing at freeway bottlenecks; part III: multi-destination flows. Transportation Research B 27, 281-313.

Papageorgiu M., 1990. Dynamic modeling, assignment and route guidance in traffic networks. Transportation Research B 24, 471-495.

Xu Y.W., Wu J.H., Florian M., Marcotte P., Zhu L.H., 1999. Advances in the continuous dynamic network loading problem. Transportation Science 33, 341-353.

Lo H.K., Szeto W.Y., 2002. A cell-based variational inequality formulation of the dynamic user optimal assignment problem. Transportation Research B 36, 421-443.