PERGAMON

Transportation Research Part B 36 (2002) 853–874

TRANSPORTATION RESEARCH PART B

www.elsevier.com/locate/trb

Shortest viable hyperpath in multimodal networks

Angélica Lozano [a,*], Giovanni Storchi [b]

[a] *Instituto de Ingeniería, Universidad Nacional Autónoma de México, Apdo. Postal 70-472, Coyoacán 04510 México DF, Mexico*

[b] *Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università di Roma -La Sapienza-, Piazzale Aldo Moro 5, 00185 Roma, Italy*

Received 7 August 2000; received in revised form 6 August 2001; accepted 10 August 2001

Abstract

In this work both the multimodal hypergraph and the viable hyperpath conceptualizations are presented. The shortest viable hyperpath problem (SVHP) in a multimodal transportation network is defined. We consider a label correcting approach to find the shortest viable hyperpath from an origin to a destination, for different values of the upper limit of modal transfers. Such hyperpaths compose a Pareto-optimal set, from where the user could choose the "best" hyperpath according to personal preferences with respect to the expected travel time and the upper limit of modal transfers. An application example on a multimodal network is presented.

1. Introduction

A path from a given point to a specified point over a multimodal network could have an "illogical" sequence of the used modes. A path of this type cannot really be travelled by the user. The path, which respects a set of constraints on the sequence of the used modes, is a *viable* path. A multimodal path can or cannot be a viable path, but a monomodal path is always viable.

In this paper, shortest hyperpaths on multimodal networks are studied, where hyperpaths are constrained to be viable, i.e. the paths composing the hyperpaths are viable. Hence, all of the paths composing a hyperpath can really be travelled by the user.

Shortest hyperpaths on hypergraphs composed by only one public surface transportation mode have been extensively studied (e.g. Nguyen and Pallottino, 1986a,b; Nguyen et al., 1994, 2001; Marcotte and Nguyen, 1998). These hyperpaths are monomodal and then viable.

[*] Corresponding author. Tel.: +52-56-22-81-33; fax: +52-56-22-81-37.
E-mail addresses: alc@pumas.iingen.unam.mx (A. Lozano), giovanni.storchi@uniroma1.it (G. Storchi).

PII: S0191-2615(01)00038-8

Many studies on shortest paths in multimodal networks do not consider the existence of non-viable paths. In this context, important studies have been performed: Fernández et al. (1994) studied the shortest path on bimodal networks, introducing a node for each modal transfer; Pallottino and Scutellà (1998) considered the number of modal transfers in a path as an attribute in the multicriteria shortest path, where the network was expanded adding modal transfer arcs; Jourquine and Beuthe (1996) expanded the multimodal network adding an arc for each action which is performed for changing mode; Miller et al. (1995) and Miller and Storm (1996) created a modal transfer arc for representing each modal change and used dynamic segmentation to construct a virtual multimodal network; Ziliaskopoulos and Wardell (2000) studied time-dependent least time paths on multimodal networks (where each line had a fixed schedule and was treated as a different mode), without explicitly expanding the network; Crainic and Rousseau (1986) included a node to represent a set of operations done in a terminal, and generated a limited number of "good" itineraries to reduce costs and delays, and improve the quality of service.

Few studies eliminate "illogical" paths. Battista et al. (1996) studied the shortest paths on bimodal networks, eliminating paths with illogical sequences of the used modes. Lozano and Storchi (2001) studied the shortest paths in multimodal networks, eliminating the paths which do not respect a set of constraints on the sequence of used modes and exceed a fixed number of modal transfers.

Hyperpaths with multiple transportation modes can represent the user's alternatives for travelling in an urban network where public surface modes do not have exact schedules.

The aim of this paper is to contribute to the systematic study of the shortest viable hyperpaths on urban multimodal networks. Specifically, we study the shortest viable hyperpath problem (SVHP), whose solution is a set of viable hyperpaths with minimum expected travel time and a number of modal transfers lower than a number specified by the user. From this set, the user could choose the "best" hyperpath (i.e. strategy of travelling) according to personal preferences.

Hypergraph concepts are included in Section 2. Multimodal hypergraph and viable hyperpath are introduced in Sections 3 and 4, respectively. Specifically, in Section 4 the following elements are defined: state, transition of states and hypertransition of states, and their role in controlling the hyperpath viability is explained. The performance of modal transfers in hyperpaths is also explained. In Section 5, the shortest viable hyperpath problem, SVHP, is defined and an "ad hoc" algorithm to solve it is proposed. An example of the algorithm application and the conclusions are shown in Sections 6 and 7.

2. Hypergraph

A brief discussion of the hypergraphs from Nguyen et al. (1994) is included in this section to facilitate the subsequent presentation of the multimodal hypergraphs.

A *hypergraph or h-graph* is a pair $H = (N, E)$, where N is the set of nodes and E is the set of h-arcs. A *h-arc* $e = (t(e), h(e))$ is identified by its tail $t(e) \in N$ and its head $h(e) \subseteq N/t(e)$. If $|h(e)| = 1$, the h-arc is equivalent to an arc $e = (i, j)$.

A h-arc $e = (t(e), h(e))$ is a *support h-arc*, if h-arc $e' = (t(e), h'(e)) \in E \; \forall h'(e) \subseteq h(e)$. A h-arc e' is called a *contained* h-arc.

In a h-graph, a *path* q_{od} connecting destination d and origin o is a sequence of nodes and h-arcs: $q_{od} = (o = t(e_1), e_1, t(e_2), e_2, \ldots, e_m, d)$, where $t(e_{i+1}) \in h(e_i)$, for $i = 2, \ldots, m-1$, and $d \in h(e_m)$.

A *hyperpath* p_{od} is the minimal acyclic set of paths q_{od} such that the destination, d, is connected to every node belonging to p_{od}.

Let L be the set of lines of a public surface mode r. The h-graph corresponding to this mode is a support h-graph $H = (N, E)$ obtained from connecting several disjointed networks, one for each line of mode r. $l \in L$ is a path whose nodes represent the stops of this line.

L_i represents the set of lines with a stop at i $(i \in N)$. The set L_i may also be specified by the set of nodes corresponding to the line-stops at i. Hence, the support h-arc $e = (i, h(e))$, with $h(e) = \{j_1, j_2, \ldots, j_v\}$, represents the set of *boarding h-arcs* of the lines $\{l_1, l_2, \ldots, l_v\}$ with a stop at i.

A model of the user's behavior consists in assuming that each user has a priori chosen a subset $L'_i \subseteq L_i$. The *attractive set* L'_i is a set of lines to go from i to d, such that, at node i, the user is willing to board the first carrier of subset L'_i which arrives. Each attractive set L'_i is associated with a contained boarding h-arc $e'_i = (i, h(e'_i))$, where $h(e'_i)$ is the set of line-stops at i, of the lines belonging to L'_i.

A hyperpath from o to d (p_{od}) represents the set of possible paths for the user, where each boarding h-arc represents the attractive set for the user.

The following assumptions are made: passengers arrive randomly at every node stop, and always board the first carrier, of their attractive set, which arrives; and all lines are statistically independent and each line arrives to a node with exponential distribution (mean equal to the inverse of the frequency).

If φ_j is the frequency of line $j \in L_i$, then:

$$\Phi(e'_i) = \sum_{j \in h(e'_i)} \varphi_j,$$

$$\omega(e'_i) = 1/\Phi(e'_i),$$

$$\pi(e'_i, j) = \varphi_j / \Phi(e'_i),$$

where $\Phi(e'_i)$ is called *combined frequency* of the attractive set L'_i; $\omega(e'_i)$ represents the average waiting time at stop i, when the attractive set is L'_i; and $\pi(e'_i, j)$ denotes the probability that the first carrier arriving at stop i is of line $j \in L'_i$. The last coefficient is used to represent the conditional probability of using e (or arriving at node j) from node i.

A travel time $C(e) = c(i, j)$ and a coefficient $\pi(e, j) = 1$ are associated with each arc; while a waiting time $C(e'_i) = \omega(e'_i)$ and as many coefficients as the number of nodes in $h(e'_i)$, i.e. the coefficients $\pi(e'_i, j)$ $\forall j \in h(e'_i)$ such that

$$\sum_{j \in h(e'_i)} \pi(e'_i, j) = 1,$$

are associated with each h-arc $e'_i = (i, h(e'_i))$.

A value $V_p(i)$, which represents the expected travel time for going from i to d, is associated with each hyperpath p_{id}:

$$V_p(d) = 0$$

$$V_p(i) = \begin{cases} c(i, j) + V_p(j) & \text{if } e = (i, j); \\ \omega(e'_i) + \sum_{j \in h(e'_i)} \pi(e'_i, j) V_p(j) & \text{if } e'_i \text{ is a boarding } h\text{-arc.} \end{cases}$$

$V_p(o)$ is the expected travel time (of the whole trip) for an user who has the strategy for travelling represented by hyperpath p.

Let $V^*(j)$ $\forall j \in h(e)$, be the minimum values. An *active arc* $e = (i, j)$ is an arc such that $V^*(i) > V^*(j)$. Let $e = (i, h(e))$ be the support boarding h-arc at stop i. The optimal travel time from i to d, obtained using one of the boarding h-arcs contained in e, is defined as

$$C^*(e) = \min\{C(e', i) : h(e') \subseteq h(e)\},$$

where

$$C(e', i) = \omega(e') + \sum_{j \in h(e')} \pi(e', j) V^*(j)$$

$\forall e' = (i, h(e'))$ with $h(e') \subseteq h(e)$.

A h-arc $\tilde{e} = (i, h(\tilde{e}))$ is an *optimal boarding h-arc*, if it is the minimum subset of lines that minimize the expected travel time, i.e., if

$$|h(\tilde{e})| = \min\{|h(e')| \ \forall e' = (i, h(e')) : h(e') \subseteq h(e), C(e', i) = C^*(e)\}.$$

If $\tilde{e}$ also satisfies $V^*(i) > C^*(e) - \omega(\tilde{e})$, then it is called *active boarding h-arc*. Hence, for an active boarding h-arc, $V^*(i) > V^*(j)$ $\forall j \in h(\tilde{e})$.

An *active hyperpath* is a hyperpath composed of only active arcs and active boarding h-arcs. An *active h-graph* is composed of a set of active hyperpaths. An active h-graph is a support h-graph.

3. Multimodal hypergraph

Some transportation modes can be represented by hypergraphs. If a multimodal transportation system includes at least one of these modes can be represented by a multimodal h-graph.

Definition 1. A *multimodal h-graph* is a triplet $H = (N, E, M)$ where N is the set of nodes, E is the set of h-arcs and M is the set of modes associated with the h-arcs.

A monomodal mode-r [1] h-graph, $r \in M$, is represented by $H_r = (N_r, E_r)$, where N_r and E_r are the sets of nodes and h-arcs, respectively.

Only one mode-r is associated with each h-arc of a multimodal h-graph. To obtain it, modal transfer arcs, i.e. arcs with the modal transfer "mode" associated, are added. Hence, an adjacent arc of a mode-r arc is either a mode-r arc or a modal transfer arc.

Nine modes are considered here for urban networks. The modes are classified in six subsets according to their characteristics. The subset description is shown in Table 1. The mode of a h-arc is an element of the set $M = \{M1, M2, M3, M4, M5, M6\}$.

Public rail modes (subset $M1$) are the modes such that at each stop, the user could take only one line with fixed schedule. Times of waiting and changing line are included in the travel time.

A network composed of only one public surface mode, whose frequencies are known by the user but the exact schedules are not, and where at each stop the user is willing to board the first

[1] *mode-r* is used instead of mode r.

Table 1
Mode classification

Subset	Description
$M1$	Public rail modes: metro and train
$M2$	Public surface modes: bus, collective taxi and taxi
$M3$	Private modes with parking needs: car and motorcycle
$M4$	Private modes without parking needs: bicycle
$M5$	Walking mode
$M6$	Modal transfer

carrier of the attractive set of lines, can be represented by a *h*-graph. Each mode of subset $M2$ can be represented by a *h*-graph.

The modes in $M2$ could have many lines, the bus and collective taxi have a fixed number of lines, and the taxi has the lines the user needs. The user could take mode taxi for going anywhere, but in practice takes it to go either to a point where the journey can continue via another mode or to the destination.

Although $M6$ is not a mode, it is included in M to obtain a multimodal *h*-graph as defined.

Therefore, a mode-r belonging to subsets $M2$ or $M6$ is represented by a monomodal mode-r graph; and a mode-r belonging to $M2$ is represented by a monomodal mode-r *h*-graph. These graphs and *h*-graphs are connected amongst themselves by mode $M6$ arcs, composing then a multimodal *h*-graph.

$M6$ can represent different actions depending on the modes involved in the change. The change of mode is a process which could include the following actions:

(i) leaving a line of the current mode;
(ii) leaving the current mode;
(iii) approaching the new mode-r;
(iv) waiting for mode-r and
(v) taking mode-r.

A multimodal *h*-graph is composed of three types of *h*-arcs: boarding *h*-arcs, travel arcs and modal transfer arcs.

A boarding *h*-arc represents actions (iv) and (v), and exists only for the modes of $M2$ (i.e. the modes which can be represented by *h*-graphs).

A *travel arc* of mode-r represents either the action of travelling by mode-r or the actions of travelling by a line of mode-r and leaving such a line (action (i)), for $r \in (M1 \cup M2)$. The act of leaving mode-r, $r \in (M1 \cup M2)$, means not taking any mode-r line immediately. This action is not included in a travel arc.

A *modal transfer arc* could represent either actions (ii) and (iii) or actions (ii), (iii) and (v), depending on the modes involved in the change.

Travel arcs and modal transfer arcs have an associated time, unlike boarding *h*-arcs that have associated the frequencies of the lines that can be boarded.

A change from a mode-r_1 to a mode-r_2 (with $r_1 \neq r_2$) has one of the following representations:

- If $r_1, r_2 \notin M2$, the change of mode includes actions (ii), (iii) and (v), and then can be represented only by a modal transfer arc.
- If $r_1 \notin M2$ and $r_2 \in M2$, the change of mode includes actions (ii), (iii), (iv) and (v), and then can be represented by a modal transfer arc (actions (ii) and (iii)) and a boarding *h*-arc (actions (iv) and (v)).
- If $r_1 \in M2$ and $r_2 \notin M2$, the change of mode includes actions (i), (ii), (iii) and (v), and then can be represented by a travel arc (action (i)) and a modal transfer arc (actions (ii), (iii) and (v)).
- If $r_1, r_2 \in M2$, the change of mode includes actions (i), (ii), (iii), (iv) and (v), and then can be represented by a travel arc (action (i)), a modal transfer arc (actions (ii) and (iii) and a boarding *h*-arc (actions (iv) and (v)). But if the nodes involved in the change coincide (for example, when a bus stop coincides with a collective taxi stop), action (iii) does not exist and then the modal transfer arc represents only action (ii).

Therefore, a modal transfer arc is not enough to represent a modal change where a mode $M2$ is involved.

If instead of changing the mode, the user changes the line keeping the mode, then a *transboard* is made. A transboard is a change from one line to another, both of the same mode-*r*.

A multimodal hypergraph can or cannot include a representation of transboards. In this study, the number of transboards an user can do is not restricted. Then in the following sections a multimodal *h*-graph without representation of transboard is used. Multimodal *h*-graphs with representation of transboards are not considered here due to their complexity.

4. Viable hyperpaths

If several transportation modes are available, the "best" multimodal hyperpath (p_{od}) that can be travelled is sought.

4.1. Viable hyperpaths in a multimodal h-graph

Some definitions of hyperpaths composed of one or multiple modes are presented as follows:

Definition 2. A *monomodal mode-r hyperpath* is a hyperpath composed of nodes which belong to the same N_r set, $r \in M2$.

Definition 3. A *multimodal hyperpath* is a hyperpath composed of nodes which belong to several N_r sets, $r \in M$.

Definition 4. A mode-*r* is a *constrained mode* if only one maximal mode-*r*-subpath [2] is allowed in a path.

[2] A *maximal mode-r-subpath* is a subpath formed by nodes belonging to N_r, where the initial and the ending nodes are adjacent to nodes belonging to other sets (Lozano and Storchi, 2001).

A constrained mode can be used only once in a viable path. Let $M_v, M_v \subseteq M$, be the set of constrained modes.

Definition 5. A multimodal hyperpath is *viable* if the paths composing the hyperpath do not include more than one maximal mode-r-subpath for each $r \in M_v$. That is, the paths of a viable multimodal hyperpath do not use a constrained mode more than once.

A mode-r, $r \in M2$, has the following two important characteristics:

(a) a mode-r network is represented by a mode-r h-graph;
(b) a mode-r is not constrained, i.e. any path of a hyperpath can include more than one maximal mode-r-subpath.

The generation of a new viable hyperpath through the concatenation [3] can be performed in two different ways:

1. Let $e = (i, j)$ be a mode-r arc and let p_{jd} be a viable hyperpath from j to d. The concatenation of p_{jd} with arc e is a viable hyperpath p_{id} if, the concatenation of arc e with each viable path (q_{jd}) that composes the hyperpath p_{jd} generates a viable path q_{id}.
2. Let $e'_i = (i, h(e'_i))$ be a boarding h-arc (i.e. a mode-r h-arc, $r \in M2$), and let p_{jd} $\forall j \in h(e'_i)$ be the viable hyperpaths from nodes j to d. The concatenation of the viable hyperpaths (p_{jd}) with h-arc e'_i is always a viable hyperpath p_{id}.

Hence, the verification of the viability of the new hyperpath p_{id} must be performed only if the h-arc is not of mode-r, $r \in M2$ (i.e. if the h-arc is an arc). Therefore, for an arc (i, j) of mode-r, $r \notin M2$, it must be verified that the concatenation of the arc with each path composing the viable hyperpath p_{jd} is a viable path q_{id}.

4.2. States of viable hyperpaths

The *state s* of a viable path is a key to indicate an admissible composition of the modes on the viable path (Lozano and Storchi, 2001). A path with an associated state is a viable path which has a specific sequence of the used modes. The state is used to check viability of the paths obtained from the concatenation.

Since a viable hyperpath p_{id} is composed of a set of viable paths q_{id}, each one with an associated state, then a state can also be associated with the viable hyperpath. This state indicates the specific sequence of the used modes in all of the paths composing the hyperpath.

It is assumed that, in a trip, the user can only use: one mode of $M1$, three modes of $M2$ and one mode of each one subset $M3$, $M4$ and $M5$. All of these modes are constrained, except modes of $M2$ and $M5$. Hence, there are few possible sequences of used modes in a viable path. Only 20 states

[3] *Concatenation* of the paths $P1$ and $P2$, denoted by $(P1 \circ P2)$, is a path P formed by $P1$ and $P2$ such that the destination of $P1$ is the origin of $P2$ (Lozano and Storchi, 2001).

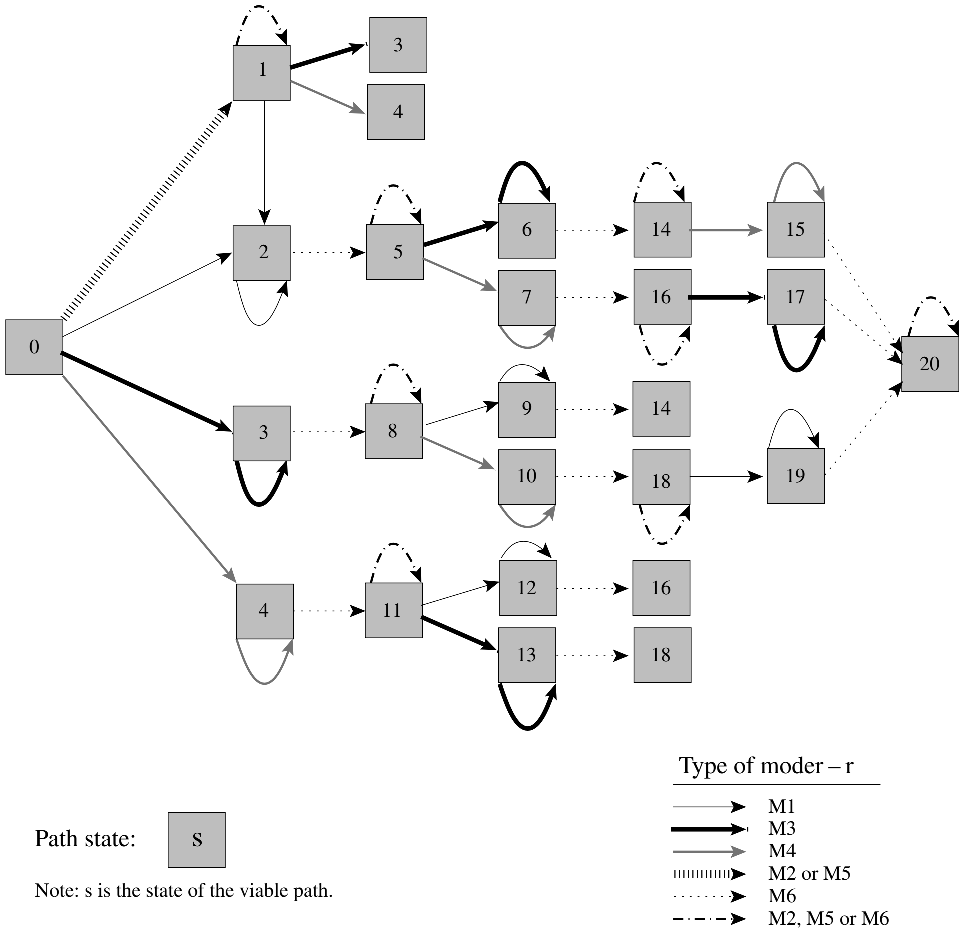

Fig. 1. Transition of states for viable paths.

related to such sequences can be associated with viable paths. The states are shown in Fig. 1. In case of having less modes, the number of states can decrease.

In Fig. 1, a node represents a state associated with a viable path, and an *Mi* arc represents a journey via a mode-*r* of subset *Mi*. A trip starts at the origin without an associated state (node 0).

For example (see Fig. 1), if a trip stars via a mode $M1$ ($M1$ arc), a path of state 2 is obtained (node 2); while the journey continues via the mode $M1$, paths of state 2 are obtained, but if such mode is left ($M6$ arc), a path of state 5 is obtained (node 5). Then the journey can continue via modes $M2, M5$ or $M6$ ($M2$, $M5$ or $M6$ arcs) obtaining paths of the same state 5; via mode $M3$ ($M3$ arc) obtaining a path of state 6 (node 6); or via mode $M4$ ($M4$ arc) obtaining a path of state 7 (node 7); but it cannot continue via mode $M1$ because $M1$ is a constrained mode ($M1 \in M_v$) which has been used and left before.

A path of state 6 is a path which has used and left mode $M1$ ($\in M_v$) and probably some unconstrained modes, and has used but has not left mode $M3$ ($\in M_v$). It is possible to continue this path as follows: (a) via mode $M3$, obtaining another path of state 6; (b) via some unconstrained mode, getting a path of state 14; or (c) via mode $M4$, obtaining a path of state 15; but it is not possible to continue the journey via the constrained mode $M1$.

In Fig. 1, an arc of mode-r from state s_1 to state s represents a journey via mode-r starting with a path of state s_1 and ending with a path of state s. The existence of such an arc indicates that the transition [4] from state s_1 to state s through an arc of mode-r is possible. When $s_1 = s$, the arc is a loop.

Given that the concatenation of a set of hyperpaths with a h-arc can be performed in two different ways, depending on the h-arc characteristics, the assignment of a state to the new hyperpath, obtained from the concatenation, can also be performed in two differentways:

1. If $e = (i, j)$ is a mode-r arc and hyperpath p_{jd} has an associated state s_x, then the transition of states shown in Fig. 1 is applied. If the corresponding transition is possible, then the new hyperpath p_{id} is viable and a state s is assigned to it.
2. If e'_i is a mode-r h-arc ($r \in M2$) and each viable hyperpath p_{jd} $\forall j \in h(e'_i)$ has an associated state, then the hyperpath generated from their concatenation is always viable and has an associated state. The state is assigned based on the hypertransition of states, which is described in the following section.

4.3. Hypertransition of states

Let p_{id} be the concatenation of a h-arc e'_i with the set of viable hyperpaths p_{jd}. Given that the h-arc is of mode-r, $r \in M2$, the hyperpaths p_{jd} have to leave node j using a travel arc of the same mode-r. Therefore, the states which can be associated with the hyperpath p_{id} are the states which in Fig. 1 have an $M2$ arc (i.e. states 1, 5, 8, 11, 14, 16, 18 and 20).

Let S be the set of states. The *hypertransition* is a binary operation on S. This operation, denoted by $*$, satisfies the closure, commutative and associative properties, that is: $s_z * s_x \in S$ $\forall s_z, s_x \in S$; $s_z * s_x = s_x * s_z$ $\forall s_z, s_x \in S$; and $(s_z * s_y) * s_x = s_z * (s_y * s_x)$ $\forall s_z, s_y, s_x \in S$.

Definition 6. The hypertransition from states s_z and s_x to state s, $s_z * s_x = s$ is the specification of a state s which indicates the sequence of modes used in both the hyperpaths of states s_z and s_x.

Hence, the hypertransition of states is the transformation from states s_z and s_x to state s.

A path which has used M_v modes has a limited choice for its extension. If at least one of the paths composing a hyperpath has used and left an M_v mode, the hyperpath cannot be continued using such a mode. Then, a hyperpath, whose paths have used M_v modes, has a limited possibility of extension. A preference relationship between states is established, on the basis of the possibility of extension of the corresponding paths. The Preference Theorem [5] establishes whether a state s_u is preferred over a state $s_l, s_u R s_l$ (where R is a strict preorder). Using the Preference Theorem,

[4] The *transition* from states s_1 to s through an arc of mode-r ($r \in M$) is the action of obtaining a state s path by means of concatenation of a state s_1 path with an arc of mode-r (Lozano and Storchi, 2001).

[5] *Preference Theorem*: State s_u is preferred over state s_l, $s_u R s_l$, if the following sentences are valid: (a) the constrained modes used in state s_u path, are a subset of the set of constrained modes used in state s_l path; (b) if state s_l path is using a constrained mode, then state s_u path is using this mode or has never used it (Lozano and Storchi, 2001).

Table 2
Sets of non-preferred states compared with each state s ($NPS(s)$ sets)

State s	$NPS(s)$
1	$1, 2, 3, \ldots, 20$
5	5, 6, 7, 14, 15, 16, 17, 20
8	8, 9, 10, 14, 15, 18, 19, 20
11	11, 12, 13, 16, 17, 18, 19, 20
14	14, 15, 20
16	16, 17, 20
18	18, 19, 20
20	20

we define $NPS(s_u)$ as the set of states which are not preferred compared with state s_u; i.e. $NPS(s_u) = \{s_v : s_u R s_v; s_u, s_v \in S\}$.

Table 2 shows the $NPS(s)$ set for each possible state s. The elements of an $NPS(s)$ are shown in Fig. 1 as all the possible successive nodes of node s, including itself.

From the states which appear in both sets $NPS(s_z)$ and $NPS(s_x)$, the state preferred over the others is chosen. This is the state s, which results from the hypertransition $s_z asts_x = s$. State s appears in the intersection of row s_z and column s_x of Table 3.

An example of hypertransition is the following. Suppose that $e' = (i, h(e'))$ is a h-arc with $h(e') = \{j_1, j_2, j_3\}$, and s_z, s_x and s_y are the states of hyperpaths p_{j_1d}, p_{j_2d} and p_{j_3d}, respectively. Also, suppose that p_{j_1d} is the first hyperpath concatenated with the h-arc. The resulting state of this first hypertransition is s_z (the state of p_{j_1d}), which will associate with p_{id}. When p_{j_2d} is also concatenated with the h-arc, the hypertransition goes from states s_z (the state of the current p_{id}) and s_x (the state of p_{j_2d}) to state s, which will associate with the new p_{id}. When p_{j_3d} finally is concatenated with the h-arc, the hypertransition goes from states s and s_y to another state which will associate with the new p_{id}.

Therefore, the concatenation of a set of hyperpaths with a h-arc, where $|h(e')| > 2$, is performed in several steps, concatenating only one hyperpath each time until exhausting the nodes of the h-arc head (i.e., the states of the hyperpaths).

Table 3
Hypertransition of states

s_z	s_x							
	1	5	8	11	14	16	18	20
1	1	5	8	11	14	16	18	20
5		5	14	16	14	16	20	20
8			8	18	14	20	18	20
11				11	20	16	18	20
14					14	20	20	20
16						16	20	20
18							18	20
20								20

4.4. Modal transfers in a viable hyperpath

If arc (i, j) is concatenated with hyperpath p_{jd}, then it is possible to determine the number of modal transfers in the hyperpath p_{id}. If the arc has mode $M6$, the number of modal transfers in p_{id} exceeds by one unit the number of modal transfers in p_{jd}; otherwise, the number of modal transfers in p_{id} and in p_{jd} is equal.

If h-arc e' is concatenated with the set of viable hyperpaths (p_{jd}), each one with a specific number of modal transfers, then it is not possible to determine the number of modal transfers in the hyperpath p_{id}, but it is possible to assign a number to it. This number is taken from the number of modal transfers, which were assigned to the paths forming the hyperpath p_{id}, being equal to the mean, upper or lower value.

The user assigns a personal cost to modal transfers, and does not wish to perform more than k. Hence, the user wants to avoid the risk of finding in the path a number of modal transfers greater than k. Therefore, we establish that the assigned number of modal transfers for p_{id} must be equal to the upper value of the number of modal transfers associated with hyperpaths p_{jd} $\forall j \in h(e')$.

Hence, $w(i) = \max\{w(j) : j \in h(e')\}$, where $w(i)$ indicates the upper number of modal transfers, which has been associated with some path composing p_{id}. This number is denominated *upper limit of modal transfers*.

5. Shortest viable hyperpath

In an urban area, the user searches for the ‘‘best’’ way to go from a certain origin to a specific destination through a multimodal transportation network. If the public surface modes have lines without established schedules, but the user is aware of their frequencies, then the user confronts a problem which we call *shortest viable hyperpath problem* (SVHP).

5.1. SVHP definition

The SVHP is to find the viable hyperpaths with the minimum expected travel time, where the user does not have to execute more than k modal transfers.

The SVHP solution is a set of shortest viable hyperpaths with modal transfers between 0 and k. These hyperpaths have different values associated with the following two criteria: the expected travel time and the upper limit of modal transfers. This set is a Pareto-optimal set. [6]

The generation of a Pareto-optimal set is the first stage in solving a bicriteria problem. The second stage is the subjective choice, from among the hyperpaths of this set, of the hyperpath considered the ‘‘best’’ according to the user’s personal preferences regarding both criteria.

[6] A Pareto-optimal set contains only the non-dominated solutions, such that any improvement in one single criterion could be achieved only by simultaneously damaging or reducing another criterion.

5.2. SVHP algorithm description

A procedure, based on the works of Pallottino and Scutellà (1997), Lozano and Storchi (2001) and Nguyen et al. (1994), is proposed for solving the SVHP. The procedure finds the set of hyperpaths from a certain origin to a specified destination, with the minimum expected travel time and upper limit of modal transfers between 0 and k.

The algorithm keeps the expected travel time, the upper limit of modal transfers and the state, for the hyperpaths from each visited node i to the destination.

The viable hyperpaths from node i to the destination d (p_{id}) could have different associated states. A node-state couple, $[i,s]$, represents a viable hyperpath p_{id} of state s.

Five labels that indicate the characteristics of a viable hyperpath p_{id} of state s are associated with each couple $[i,s]$. The labels are the following:

$V_s^*(i)$ = expected travel time of the current shortest hyperpath $[i,s]$.

$w_s(i)$ = upper limit of modal transfers in the current shortest hyperpath $[i,s]$.

$SA_s(i)$ = successive h-arc of node i, in the current shortest hyperpath $[i,s]$.

$ST_s(i)$ = set of states, of the shortest hyperpaths, currently in the head of h-arc $SA_s(i)$.

$lastlabel_s(i)$ = expected travel time of the last shortest hyperpath $[i,s]$ with upper limit of modal transfers lower than $w_s(i)$.

$lastlabel_s(i)$ is used to eliminate dominated hyperpaths, and $ST_s(i)$ and $SA_s(i)$ are used to reconstruct hyperpath $[i,s]$.

For each h-arc e with $|h(e)| > 1$, the following information is conserved:

$h(e)$ = set of nodes composing the head of e;

$\Phi(e)$ = frequency of lines which can be taken through e;

$C^*(e)$ = minimum expected travel time from i to d, obtained using e.

A couple $[i,s]$ is *labelled* when the expected travel time of the hyperpath p_{id} of state s has been improved using arc (i,j).

Qnow and *Qnext* sets are used to analyze labelled node-state couples, in non-decreasing order with respect to the upper limit of modal transfers of hyperpaths. If arc (i,j) is not a modal transfer, then $[i,s]$ is inserted into *Qnow*; otherwise $[i,s]$ is inserted into *Qnext*.

At each iteration only the *Qnow* elements are examined and then eliminated from *Qnow*, until *Qnow* becomes empty; then *Qnow* takes the *Qnext* elements, *Qnext* becomes empty and $(k'+1)$-iteration starts. This procedure is repeated until $k' > k$ *or Qnow* is empty.

In order to control dominance between the expected travel time and the upper limit of modal transfers, $lastlabel_{s_x}(j)$ is used. At k'-iteration if $V_{s_x}^*(j) < lastlabel_{s_x}(j)$, then the hyperpath p_{jd} of state s_x with upper limit of modal transfers equal to k' is not dominated and then $lastlabel_{s_x}(j) = V_{s_x}^*(j)$; otherwise, the hyperpath p_{jd} of state s_x is dominated and then is no longer considered.

The viability analysis of the concatenation is performed in two different ways, depending on the new h-arc characteristics. If $|h(e)| = 1$, e is an arc and then the *Arc-Concatenation* procedure is performed; otherwise, the *h-Arc-Concatenation* procedure is performed.

The *Arc-Concatenation* procedure is used to concatenate arc $e = (i,j)$ with the hyperpath p_{jd} of state s_x. This procedure has the following steps:

1. To determine, if it exists, the state s of the new viable hyperpath p_{id} and its upper limit of modal transfers.

2. To compare the expected travel time of the new hyperpath p_{id} of state s, with: the expected travel time of the current hyperpath p_{id} of state s, and the expected time of the hyperpaths p_{id} of states s_y which are preferred over state s. If the time of the new hyperpath is lower than both, the time of the current hyperpath p_{id} of state s and the time of at least one hyperpath p_{id} of any state s_y, then go to the next step; otherwise, the new hyperpath is dominated and is not generated.
3. To concatenate hyperpath $p_{jd}, j \in h(e)$, of state s_x with arc e, generating the new viable hyperpath p_{id} of state s.

The *h-Arc-Concatenation* procedure is used to concatenate a h-arc such that $|h(e)| > 1$ (i.e., a $M2$ h-arc) with a hyperpath p_{jd}, $j \in h(e)$. This concatenation is always viable, but it is executed only if the generated hyperpath is not dominated. The procedure has the following steps:

1. To determine the state s of hyperpath p_{id}, through procedure *Determine-s*. This procedure assigns a state s to a viable hyperpath p_{id}, according to the hypertransition of states.
2. To bring up to date (for the h-arc e) the following: the set of e head nodes ($h^*(e)$); the combined frequency of $e(\Phi^*(e))$; the set of states of the shortest viable hyperpaths $p_{jd}(state(e))$; the state of the shortest viable hyperpath p_{id} that uses $e(s_z(e))$; and the minimum expected travel time from i

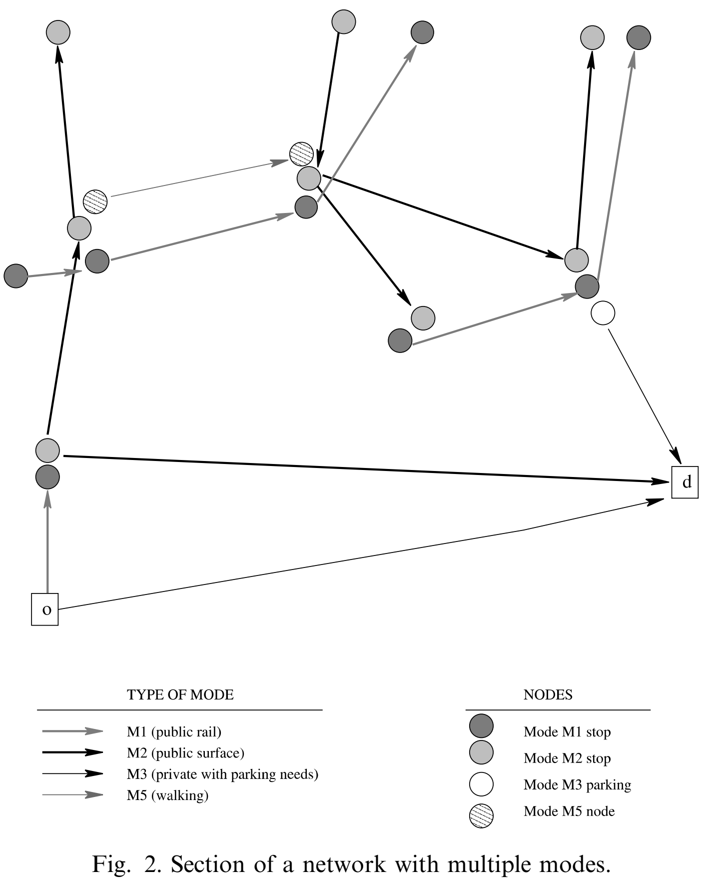

Fig. 2. Section of a network with multiple modes.

to d, using $e(C^*(e))$; when the expected travel time of the shortest hyperpath p_{jd} of state s_x is lower than the expected travel time of the shortest hyperpath p_{id} of state s (i.e. when e is active).
3. To compare $C^*(e)$ with the expected travel time of the hyperpaths p_{id} of the same state s and of other states s_y which are preferred over s. If $C^*(e)$ is lower than the expected time of the current hyperpath p_{id} of state s and the expected time of at least one of the hyperpaths p_{id} of any state s_y, then go to the next step. Otherwise the new hyperpath is dominated with respect toboth criteria and is not generated.
4. To concatenate hyperpath p_{jd} of state s_x with e, generating hyperpath p_{id} of state s. In the concatenation, $w_s(i) = w_{s_x}(j)$ because the node analysis is executed in a non-decreasing order with respect to the upper limit of modal transfers.

At the end of k-iteration, the algorithm generates the SVHP solution set. This set is composed of the viable hyperpaths from o to d, with minimum expected travel time, upper limit of modal transfers between 0 and k, and various states associated.

From the solution set only non-dominated hyperpaths with respect to both criteria (expected travel time and modal transfers upper limit) are preserved, obtaining a Pareto-optimal set.

The user chooses the "best" hyperpath according to personal preferences with respect to both criteria, from among the hyperpaths of the Pareto-optimal set.

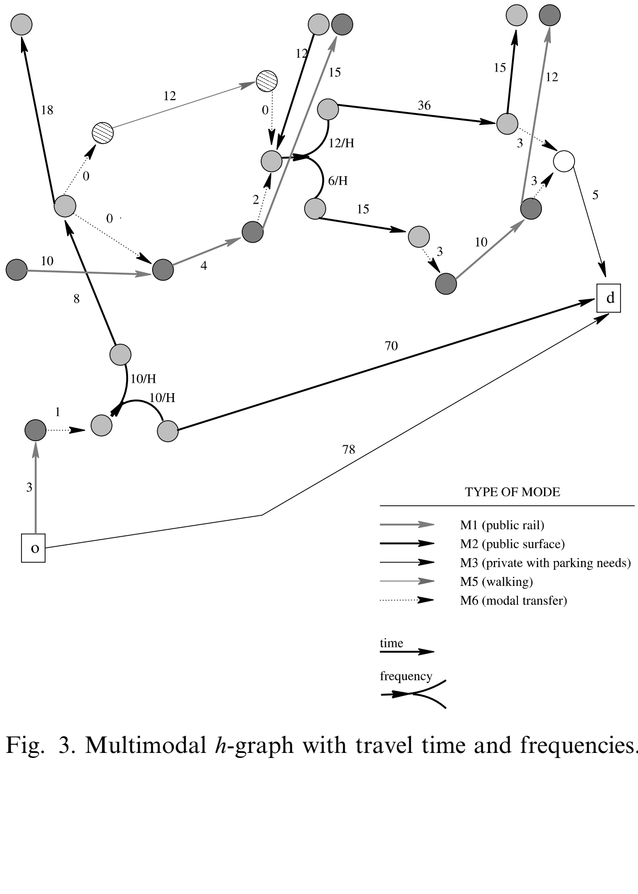

Fig. 3. Multimodal h-graph with travel time and frequencies.

6. Example

Fig. 2 shows a section of a multimodal network with four modes. There is a mode-r network for each mode, but these networks are not connected among themselves. Modal transfers are not represented. The multimodal network can be represented by a multimodal h-graph (see Fig. 3), where modal transfer arcs are included and the mode $M2$ is represented by a h-graph.

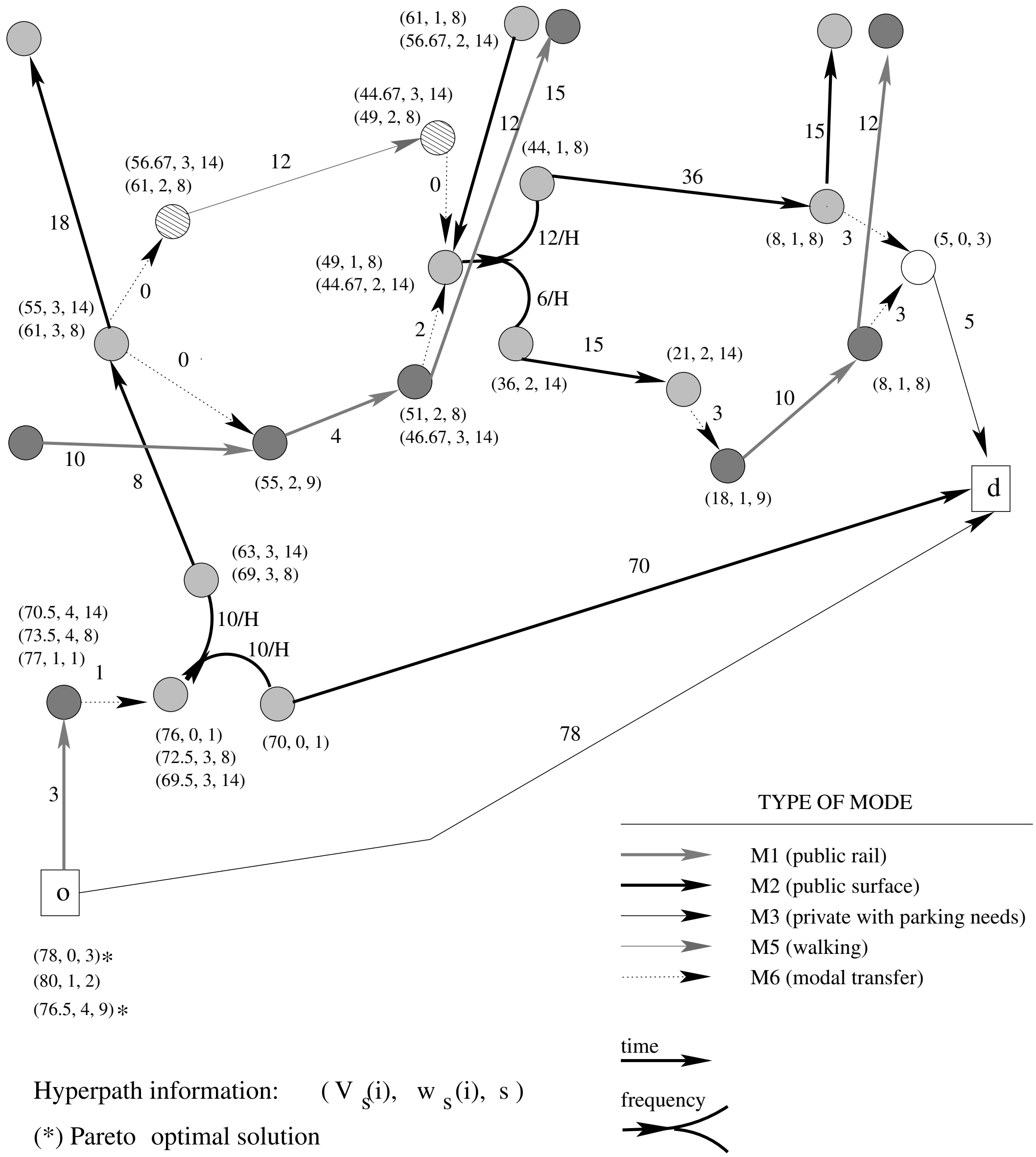

Fig. 4. Multimodal h-graph with information of the hyperpaths p_{id} of the visited nodes.

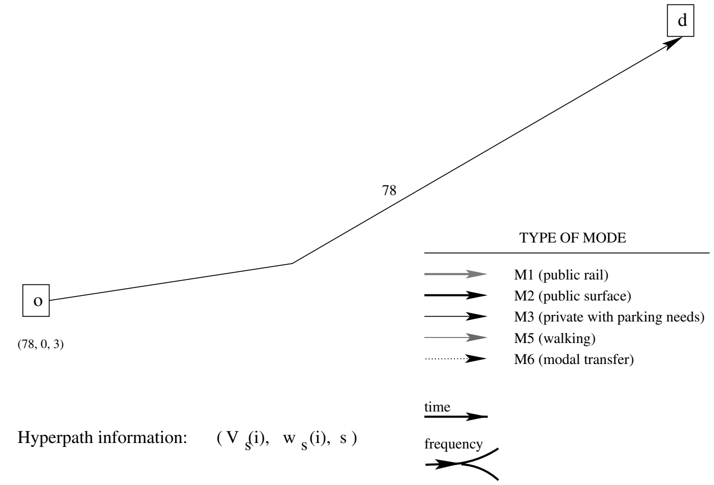

Fig. 5. Hyperpath with an expected travel time equal to 78 units, 0 modal transfers and state 3.

The results of applying the proposed algorithm on the multimodal *h*-graph are presented in Fig. 4. Stripes of three values are shown near each visited node i. These values indicate the characteristics of a shortest viable hyperpath p_{id} : $V_s(i)$ is the expected travel time of the hyperpath of state s; $w_s(i)$ indicates the upper limit of modal transfers of the hyperpath, and s is the state of the hyperpath.

There are three shortest viable hyperpaths p_{od} (from the origin o to the destination d). Since the state is no longer relevant for hyperpaths p_{od}, the hyperpath of state 2 is (with respect to both criteria) dominated by the hyperpath of state 3. Therefore, only two Pareto-optimal solutions are obtained.

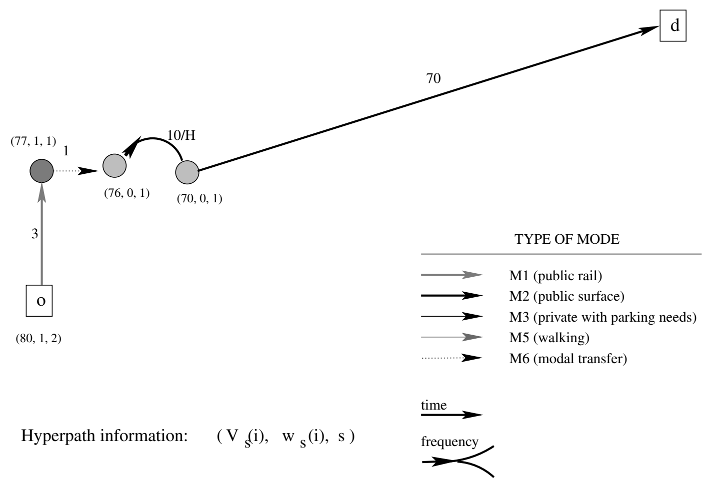

Fig. 6. Hyperpath with an expected travel time equal to 80 units, an upper limit of modal transfers equal to 1 and state 2.

The first listed hyperpath is a monomodal mode-$M3$ path, which has an expected travel time equal to 78 units, 0 modal transfers and state 3 (see Fig. 5). This hyperpath is a Pareto-optimal solution.

The second listed hyperpath has an expected travel time equal to 80 units and an upper limit of modal transfers equal to 1. This hyperpath is a bimodal path composed of two maximal mode-r subpaths: one, of the constrained mode $M1$ and the other, of mode $M2$ (see Fig. 6). This hyperpath is not a Pareto-optimal solution because it is dominated by the first hyperpath.

The last hyperpath, presented in Fig. 7, has an expected travel time equal to 76.5 units, an upper limit of modal transfers equal to 4, and state 9. This hyperpath is composed of two paths: one with five maximal mode-r subpaths of modes $M1$, $M2$, $M5$, $M2$ and $M3$ (in this order), and the other with two maximal mode-r subpaths of modes $M1$ and $M2$. The first path uses two constrained modes, while the second uses only one constrained mode. Both paths do not include more than four modal transfers. This hyperpath is a Pareto-optimal solution.

Assuming that more than $k = 4$ modal transfers are not acceptable, the "best" hyperpath must be chosen from the two hyperpaths forming the Pareto-optimal set (see Figs. 5 and 7). The decision depends on the user's personal preferences with respect to the expected travel time and the upper limit of modal transfers.

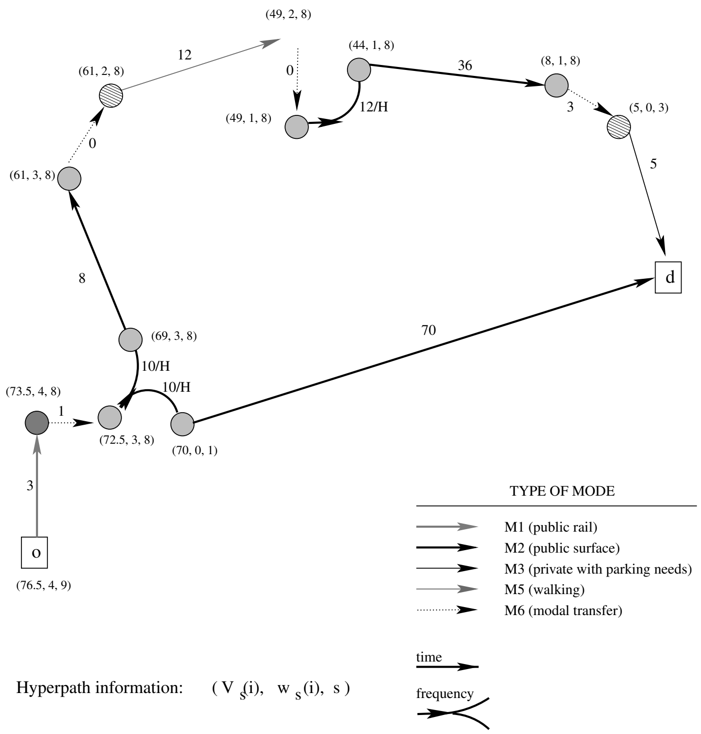

Fig. 7. Hyperpath with an expected travel time equal to 76.5 units, an upper limit of modal transfers equal to 4 and state 9.

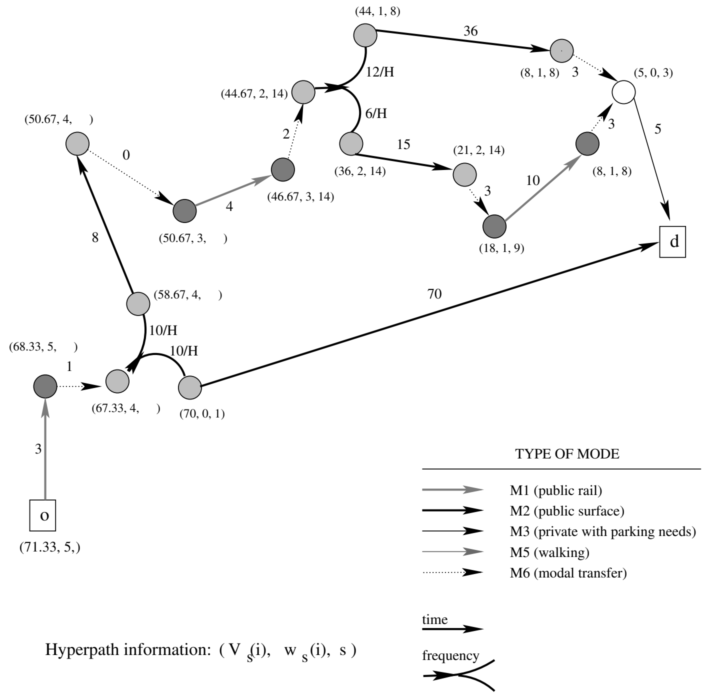

Fig. 8. Non-viable hyperpath with an expected travel time equal to 71.33 units and an upper limit of modal transfers equal to 5.

Supposing that viability would not be controlled, then the shortest (non-viable) hyperpath would have an expected travel time equal to 71.33 units and an upper limit of modal transfers equal to 5, but it would use three times the constrained mode $M1$ (see Fig. 8). Note that non-viable hyperpaths do not have an associated state.

The user, who chooses a hyperpath p_{od} (from the origin o to the destination d) with certain values $V_s(o)$ and $w_s(o)$, expects the travel last $V_s(o)$ time units and certainly non-exceeding $w_s(o)$ modal transfers.

7. Conclusions

This work contributes to the systematic study of hyperpaths on urban multimodal transportation networks. The hyperpath viability has been introduced to obtain only those hyperpaths which the user would really travel.

The proposed procedure is useful in finding a set of Pareto-optimal shortest viable hyperpaths. An upper limit of modal transfers is given by each user, in order to obtain a viable solution set. The choice of the "best" hyperpath from the solution set is a personal decision which depends on the users' preferences with respect to the expected travel time and the upper limit of modal transfers.

The procedure may be included in a traveller information system, where the user could identify optimal strategies for travelling on a multimodal network where public surface modes do not have fixed schedules.

The procedure could be included in the solution of other more complex problems on urban multimodal transportation networks, for modelling user route choice for transit-traffic assignment.

Acknowledgements

The authors wish to express their gratitude to Stefano Pallottino for his valuable comments.

Appendix A. Procedure

The main procedure of the SVHP algorithm, is:

Procedure SHORTEST-VIABLE-HYPERPATH-PROBLEM

{
for each j **in** N **do** $V^*_{s_x}(j) = \infty \ \forall s$
for each e **in** E **do**
  $\{C^*(e) = \infty; h^*(e) = \emptyset; \Phi^*(e) = 0\}$
$Qnow = \{[d, 0]\}$; $V^*_0(d) = 0$
repeat {
  while $Qnow \neq \emptyset$ {
    select $[j, s_x]$ **from** $Qnow$
    $Qnow = Qnow/\{[j, s_x]\}$
    if $(V^*_{s_x}(j) < lastlabel_{s_x}(j))$ {
      $lastlabel_{s_x}(j) = V^*_{s_x}(j)$
      for each $e \in B(j)$ **do** {
     $i = t(e)$
     if $(|h(e)| = 1)$
       call *Procedure Arc-Concatenation* (e, s_x, V^*, c, w)
     else
       call *Procedure h-Arc-Concatenation* $(e, s_x, V^*, C^*, w, h^*, \varphi)$
      }
    }
    }
    $Qnow = Qnext; Qnext = \emptyset; h = h + 1$
  } **until** $h > k$ **or** $Qnow = \emptyset$
  }

Procedure Arc-Concatenation (e, s_x, V^*, c, w)

{
  $wt = 0$; $s = 0$; $con = 0$
  switch $(mode_e)$
  case $(\in M1 :)$
    if ($s_x \neq 5$, 14, 16 **and** 20) **call** *State-M1*(s_x, s)
  break

case $(\in M3:)$
  if ($s_x \neq 8$, 14, 18 **and** 20) **call** *State*-$M3(s_x, s)$
break
case $(\in M4:)$
  if ($s_x \neq 11$, 16, 18 **and** 20) **call** *State*-$M4(s_x, s)$
break
case $(\in M6:)$
  if ($w_{s_x}(j) < k$ **and** $s_x \neq 0$) $wt = 1$; **call** *State*-$M6(s_x, s)$
break
default:
  if $(s_x = 0)$ $s = 1$
  else $s = s_x$
break
endswitch
if ($s \neq 0$ **and** $V^*_{s_x}(j) + c(e) < V^*_s(i)$) **call** *States*$(s, SM)$
else $SM = \emptyset$
while ($EOF(SM) = 0$ **and**$con \neq 1$) {
select s_y **from** SM
if $(V^*_{s_x}(j) + c(e) < V^*_{s_y}(i))$ {
  $V^*_s(i) = V^*_{s_x}(j) + c(e)$
  $w_s(i) = w_{s_x}(j) + wt$
  $SA_s(i) = e$
  $ST_s(i) = s_x$
  if ($wt = 0$ **and** $[i, s] \notin Qnow$) $Qnow = Qnow \cup \{[i, s]\}$
  if ($wt = 1$ **and** $[i, s] \notin Qnext$) $Qnext = Qnext \cup \{[i, s]\}$
  $con = 1$
}
}
}
Procedure h-Arc-Concatenation $(e, s_x, V^*, C^*, w, h^*, \varphi)$
{
  $con = 0$
  if $(\Phi^*(e) \neq 0)$ **call** *Determine-s* $(s, s_x, s_z(e))$
  else $s = s_x$
  if $(V^*_{s_x}(j) < V^*_s(i))$ {
  $\Phi^*(e) = \Phi^*(e) + \varphi_j$
  $h^*(e) = h^*(e) \cup \{j\}$
  $state(e) = state(e) \cup \{s_x\}$
  $s_z(e) = s$
  if $(\Phi^*(e) = \varphi_j) C^*(e) = 1/\varphi_j + V^*_{s_x}(j)$
  else $C^*(e) = C^*(e) - (C^*(e) - V^*_{s_x}(j))\varphi_j/\phi^*(e)$
  if $(C^*(e) < V^*_s(i))$ **call** *States*$(s; PS)$
  else $PS = \emptyset$
  while ($EOF(PS) = 0$ **and** $con \neq 1$) {

```
        select s_y from SM
        if (C*(e) < V*_{s_y}(i)) {
        V*_s(i) = C*(e)
        w_s(i) = w_{s_x}(j)
        SA_s(i) = e
        ST_s(i) = state(e)
        if ([i,s] ∉ Qnow) Qnow = Qnow ∪ {[i,s]}
        con = 1
        }
      }
      }
}
Procedure Determine-s (s, s_x, s_z(e))
{
    sum = s_z(e) + s_x
    if (s_z(e) = 1) s = s_x
    else if (s_x = 1 or (s_z(e) = s)) s = s_z(e)
      else if (sum = 19) {
        if (s_z(e) = 14 or s_x = 14) s = 14
        else s = 18
      }
      else if (sum = 26 or 29) s = 18
        else if (sum = 16, 21 or 27) s = 16
          else if (sum = 13 or 22) s = 14
            else s = 20
}
```

Procedures *State-M*1, *State-M*3, *State-M*4 and *State-M*6 determine the state s of the path, which is obtained by the concatenation of a state s_x path with an arc of modes $M1$, $M3$, $M4$ and $M6$, respectively. The procedure is based on Fig. 1.

Procedure *States* identifies the set of preferred states to s, $PS(s)$ set.

Procedure *Determine-s* determines the state s resulting from the hypertransition of states s_z and s_x, where s is the state of the new hyperpath p_{id}, s_z is the state of the current hyperpath p_{id}, and s_x is the state of the hyperpath p_{jd} which has been concatenated with the h-arc. The procedure is based on Table 3.

References

Battista, M.G., Lucertini, M., Simeone, B., 1996. Path composition and multiple choice in a bimodal transportation network. In: 7th WCTR Proceedings, Sydney.

Crainic, T.G., Rousseau, J.M., 1986. Multicommodity, multimode freight transportation: a general modeling and algorithmic framework for the service network design problem. Transportation Research B 20 (3), 225–242.

Fernández, E., De Cea, J., Florian, M., Cabrera, E., 1994. Network equilibrium models with combined modes. Transportation Science 28 (3), 182–193.

Jourquine, B., Beuthe, M., 1996. Transportation policy analysis with a geographic information system: the virtual network of freight transportation in Europe. Transportation Research C 4 (6), 359–371.

Lozano, A., Storchi, G., 2001. Shortest viable path algorithm in multimodal networks. Transportation Research A 35 (3), 225–241.

Marcotte, P., Nguyen, S., 1998. Hyperpath formulations of traffic assignment problems. In: Marcotte, P., Nguyen, S. (Eds.), Equilibrium and Advanced Transportation Modelling. Kluwer Academic Publishers, Dordrecht, pp. 175–200.

Miller, H.J., Storm, J.D., Bowen, M., 1995. GIS design for multimodal network analysis. In: GIS/LIS'95 Proceedings, pp. 750–759.

Miller, H.J., Storm, J.D., 1996. Geographic information system design for networks equilibrium-based travel demand models. Transportation Research C 4 (6), 373–389.

Nguyen, S., Pallottino, S., 1986a. Hyperpaths and shortest hyperpaths. In: Simeone, B. (Ed.), Combinatorial Optimization. Lecture Notes in Mathematics. Springer, Berlin, pp. 258–271.

Nguyen, S., Pallottino, S., 1986b. Assegnamento dei passaggeri ad un sistema di linee urbane: determinazione degli ipercammini minimi. Ricerca Operativa 38, 29–74.

Nguyen, S., Pallottino, S., Malucelli, F., 1994. A modeling framework for the passenger assignment on a transportation network with time-tables. crt-94-47 University of Montreal, Quebec, CA.

Nguyen, S., Pallottino, S., Malucelli, F., 2001. A modeling framework for the passenger assignment on a transportation network with time-tables. Transportation Science 35, 238–249.

Pallottino, S., Scutellà, M.G., 1998. Shortest path algorithms in transportation models: classical and innovative aspects. In: Marcotte, P., Nguyen, S. (Eds.), Equilibrium and Advanced Transportation Modelling. Kluwer Academic Publishers, Dordrecht, pp. 240–282.

Ziliaskopoulos, A., Wardell, W., 2000. An intermodal optimum path algorithm for multimodal networks with dynamic arc travel times and switching delays. European Journal of Operational Research 125, 486–502.