Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

Routing of freeway traffic - A discrete-time state space model and routing problems

J.H. van Schuppen

Department of Operations Reasearch, Statistics, and System Theory

BS-R9232 1992

Routing of Freeway Traffic - A Discrete-Time State Space Model and Routing Problems *

J.H. van Schuppen
CWI, P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

Abstract
Freeways were built to provide traffic and transportation means. In the technologically developed societies of Europe, the U.S.A., and Japan there is a growing demand for more traffic and transportation capacity. Because building more freeways is no longer acceptable the attention of governments and traffic engineers is directed to improving the utilisation of the existing freeway network by providing information and directions to drivers. Control and signalling systems have been installed on freeways in several countries. Feasibility studies are currently performed of routing guidance in which communication with drivers takes place by display of variable direction signs or by other communication channels. This paper is an attempt to formulate an approach to routing of freeway traffic flow. A state space model is proposed for traffic flow in a freeway network. The model consists of a control system in which the state variables are densities and average speeds in freeway sections. The routing problem is to synthesize a routing law that specifies the directions such that the travel cost of network users is minimized. Approaches to solve the routing problem by control theory are discussed.

1991 Mathematics Subject Classification: 93C10, 93E20.

Keywords & Phrases: freeway network, routing, control.

1 Introduction

The purpose of this paper is to contribute to the problem of routing freeway traffic by proposing a state space model and by formulating routing problems.

Freeways were built to provide fast traffic and transportation means. In urban areas of the technologically developed societies of Europe, the U.S.A., and Japan, congestion of freeways is becoming a major obstacle to traffic flow. The solution of building more freeways is politically and environmentally no longer accepted in many countries. Attention is therefore concentrated on improving the efficiency of the existing freeway network by providing information and directions to drivers. This approach is supported by such governmental programs as the program DRIVE of the Commission of the European Communities and as the program Intelligent Vehicle/Highway System (IVHS) of the Federal Highway Administration in the U.S.A.

Freeway signalling and control systems have been installed in several countries. In The Netherlands such a system has been installed over the last 10 years in the densely populated areas around the cities of Amsterdam, Rotterdam, and Utrecht. The tasks of this system are to collect information on the state of traffic, to warn drivers on traffic conditions downstream, to improve traffic flow, and to aid police and maintenance crews. Currently studies are conducted for route guidance. In the future variable direction signs may be installed at freeway intersections on part of the freeway

*Report BS-R9232, ISSN 0924-0659, CWI, P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

network. Depending on the state of traffic in the network, drivers are advised to use the regular route or one of several alternate routes. Route guidance is considered for freeway networks in Europe. Pilot projects are planned for the mid nineties and widespread application by variable direction signs is anticipated for the year 2010. There is a need to evaluate the usefulness of route guidance. Such a feasibility study must be based on a mathematical model for traffic flow in a network and on routing algorithms. The problem of this paper is thus motivated.

The problem considered in this paper is to model traffic flow in a freeway network and to synthesize routing algorithms. The *objectives* of routing freeway traffic flow are:

1. To make the travel cost of network users as small as possible.

2. To provide fairness to all users of the network.

The effect of routing may be that congestion is prevented or postponed, and that the network is used efficiently. However, congestion and efficiency will also depend on the network's infrastructure and on traffic demand, which feedback by variable direction signs cannot affect. The *time scale* of updating variable routing directions is, say, once in every 5 to 10 minutes. The *information span* of the routing system is that information on the state of traffic in the network must be available at one central location. Aggregation of this information is necessary. The *control actuators* of the routing system are primarily variable direction signs at freeway intersections.

The approach taken to the modelling problem of freeway traffic flow is to adapt the model developed and validated by S.A. Smulders in [14, 15, 16, 17, 18, 19]. This model is based on a flow model originally proposed by M.J. Lighthill and G.B. Whitham [7], and modified by H.J. Payne [12, 13]. The model is reasonably realistic, but is in need of additional validation for conditions near congestion. The approach to the routing problem is to use control theory.

The differences between the routing problem considered in this paper and that of the existing literature are mentioned next. The approach to routing by optimization techniques is not useful for the routing problem of this paper. In this approach one is given the average flows from each origin to each destination summarized in the *origin-destination matrix* (OD-matrix) and the flow characteristics of each link of the network. One is then asked to select a routing scheme for the flows such that some criterion is optimized. This approach does not make use of the actual state of traffic such as traffic density and average speed. The dynamics of traffic flow in a network are essential for the dynamic routing problem. In addition the approach ignores the time dependence of arrival intensities especially in peak hour traffic when route guidance is mainly needed.

Routing in communication and computer networks is not useful for routing in freeway network because its dynamics are faster. The states in a communication network are the number of messages at nodes while in a freeway network it are the number of cars in the links. In a communication network centralized control is hardly feasible while in a freeway network centralized control is feasible. An entry to the literature on routing in communication networks is [1, 21]. The model for flow in a freeway traffic network proposed by M. Papageorgiou [11] is closely related to the model proposed here. The model of this paper differs from that of [11] in that a more detailed state space description is used and a measurement equation proposed. In addition the approach to the routing problem is more detailed.

A brief description of the content of the paper follows. The next section contains the description of the model. Section 3 provides a brief discussion on the filtering and estimation problem. The routing problem is analysed in section 4.

The author of this paper acknowledges useful discussions on the subject of this paper with P. Varaiya and J. Walrand. He thanks M. Papageorgiou for making available to him a preprint of [11].

2 A state space model for traffic flow in a freeway network

In this section a state space model for traffic flow in a freeway network will be presented. The objectives of the model are that it can be used for online routing of freeway traffic and that it is not too complex. Because a rerouting advice is in general issued only if there is congestion or if such a traffic condition is anticipated, the model should be realistic in describing the occurence and buildup of congestion. Because traffic intensities fluctuate during peak hours when routing is mainly needed, a model that incorporates traffic densities seems quite appropriate.

Example 2.1 By way of example a simple urban freeway network will be treated consisting of a rectangle or ring around a city with four links to neighbouring cities, see figure 1. This network has the advantage that for every origin-destination pair there are at most two routes through the network.

Since the figure provides only limited information, a brief description of the network follows. The nodes 1,2,3,4 represent starting and termination nodes of the network on links to other cities. The nodes 9,10,11,12 represent entries and exits. The nodes 5,6,7,8 represent freeway intersections at which there are neither exits nor entries. Freeway traffic may flow from node 1 via intersection 5 to node 9, or to node 12. There is also the flow from node 9 via intersection 5 to node 1 or to node 12. The flow form node 1 to node 3 has two possible routes, one via the intersections 5, 6, and 7 and another via the intersections 5, 8, and 7.

2.1 A model of a freeway network

Consider a freeway network. It is assumed that a control and signalling system is installed on this network. Such a system has a detector station at about every 500 m. of the freeway. At the location of this detector station there are two detection loops in every lane. The information provided by these loops are passages times and passages speeds of vehicles. It will also be assumed that detector loops are installed on every entry, exit, and branch of a freeway intersection of the freeway network. This assumption may not be realistic but for this investigation it will be adopted. A section of the network is a stretch of freeway between two adjacent detector stations. It follows from these assumptions that every entry, exit, or intersection is located at the beginning or end of a section.

Consider a model of a freeway network consisting of an ordered graph. The nodes of the graph correspond to detector stations. Let $I \subset Z_+$ denote the index set of the nodes. At a node there may be an entry, exit, intersection, or only the joint of two sections. Let $I_O \subset I$ denote the index set of the nodes at which there are entries or origins and starting points of the network. Let $I_D \subset I$ denote the index set of the nodes at which there are exits or destinations or terminating points of the network. Let $I_I \subset I$ denote the index set of the nodes at which there are freeway intersections. For example 2.1 these sets are

$$I = \{1, 2, \ldots, 12\}, I_0 = I_D = \{1, 2, 3, 4, 9, 10, 11, 12\}, I_I = \{5, 6, 7, 8\}.$$

A *section* of the network model is an ordered pair $(l, m) \in I \times I$ such that there is a path from node l to node m not passing any other node. The set of network sections is denoted by $S \subset (I \times I)$. A section of the network model corresponds to a freeway section. For example 2.1 the set of sections consists of

$$S = \{(1, 5), (5, 1), (5, 9), (9, 5), \ldots\}.$$

An *origin-destination pair* (OD-pair) is an ordered tuple $(i, j) \in I_O \times I_D$ where $i \in I_O$ is a node with an entry or origin and $j \in I_D$ is a node with an exit or destination. Let $OD \subset (I_O \times I_D)$ be the set of OD-pairs. For example 2.1 this set is given by

$$OD = \{(1, 2), (1, 3), (1, 4), (1, 9), (1, 10), (1, 11), (1, 12), \ldots\}.$$

A *route* of the OD-pair (i,j) is a path from node i to node j. A route is specified by a chain of sections. In general an OD-pair may admit several routes. For $(i,j) \in OD$ let $K(i,j) \subset Z_+$ be the index set of distinct possible routes and let

$$R(i,j) = \{r((i,j),k), k \in K(i,j)\}$$

denote the corresponding set of routes. A route is described as a chain of sections

$$r((i,j),k) = \{(i,i_2),(i_2,i_3),\ldots,(i_s,j)\}.$$

Consider example 2.1. For $(1,3) \in OD$ there are only two routes specified by $K(1,3) = \{1,2\}$,

$$r((1,3),1) = \{(1,5),(5,9),(9,6),(6,10),(10,7),(7,3)\}, \quad (1)$$
$$r((1,3),1) = \{(1,5),(5,12),(12,8),(8,11),(11,7),(7,3)\}. \quad (2)$$

2.2 A model for freeway traffic flow

H.J. Payne's model for freeway traffic flow will be used. This model is a modification of one introduced by M.J. Lighthill and G.B. Whitham [7]. The latter model has only traffic density as a state variable. Payne extended this model with the state variable of average speed. There is a continuous-time version [12] and a discrete-time version [13] of this model. Changes to Payne's model were proposed by M. Papageorgiou [10], by S.A. Smulders [16] whose model was validated for a freeway in The Netherlands, and by P. Varaiya and co-workers [6]. Below yet another version of Payne's model is defined that is close to that of [16].

In the first phase of this investigation a continuous-time model was proposed [20]. The discrete-time model proposed below is close to that of A. Messmer and M. Papageorgiou [8] and that of Papageorgiou [11]. The differences are pointed out below.

2.3 The state variables of the model

Consider then a freeway network. As argued above, a model of such a network consists of sections. The length of a section will usually be about 500 m. The model proposed below is realistic only with relatively short sections of a length that is at most 500 to 1000 m.

According to Payne's model, the state variables of freeway traffic flow are density and average speed of each section. Because the purpose of the model is routing, the traffic density per section must be separated out per origin-destination pair and per route. The model of this paper therefore differs from that of Papageorgiou [11]. In the latter model the traffic density is distinguished only on destination. In [8] it seems that fixed fractions of the density are used for each origin. Considering the dynamics of traffic flow in a freeway network, it seems necessary to consider a model in which the density per origin-destination pair and per route are distinguished. The route is important in this also. There is a report that the model of [11] leads to the phenomenon that in a particular situation freeway traffic flows back in the direction it came from. This may occur because only local directions are used. If a route is prescribed this phenomenon cannot occur. Later in the paper the model will be extended so that vehicles may change their route.

Denote the density of OD-pair $(i,j) \in OD$, of route number $k \in K(i,j)$, and of section $(l,m) \in S$ by $x((i,j),k,(l,m),.) : T \to R_+$ in vehicles per km per lane, veh/km.lane. Denote the density in section $(l,m) \in S$ by $x((l,m),.) : T \to R_+$

$$x((l,m),t) = \sum x((i,j),k,(l,m),t)$$

where the sum is over all $(i,j) \in OD$ and all routes $k \in K(i,j)$ that use section (l,m).

Denote the branching fraction for OD-pair (i,j), for route number $k \in K(i,j)$, for traffic flowing in section $(l,m) \in S$ to intersection $m \in I_I$ according to route k by

$$p((i,j),k,(l,m),t) = \frac{x((i,j),k,(l,m),t)}{x((l,m),t)} \tag{3}$$

This fraction represents the chance that a vehicle travelling in section (l, m) at $t \in T$ uses $(i, j) \in OD$ and route k.

Denote the average speed of all cars in section (l, m) by $v((l,m),.) : T \rightarrow R_+$.

The flow of section $(l, m) \in S$ is defined as $q((l,m),.) : T \rightarrow R_+$

$$q((l,m),t) = lane(l,m)x((l,m),t)v((l,m),t) \quad veh/h,$$

where $lane\ (l, m)$ is the number of lanes of the section.

Consider two adjacent sections, say $(l, m), (m, s) \in S$ such that at node $m \in I$ there is no entry, no exit, and no intersection. The transition flow at node m is defined by

$$\alpha q((l,m),t) + (1-\alpha)q((m,s),t) \quad veh/h. \tag{4}$$

Here $\alpha \in [0, 1]$ is a weighting factor. The transition flow is thus a convex combination of the flows in both sections. For a continuous-time model it was estimated in [16] that the value $\alpha = 0.85$ is appropriate.

2.4 The state transitions

The transitions of the state variables, density and average speed per section, are described next. An example is presented in 2.2. The transition of the density is based on a continuity equation

$$\begin{aligned} x((i,j),k,(l,m),t+1) &= x((i,j),k,(l,m),t) \\ &+ \frac{\Delta t}{Le(l,m)lane(l,m)} \left[q_{in}((i,j),k,(l,m),t) - q_{out}((i,j),k,(l,m),t) I_{(x((i,j),k,(l,m),t)>0)}\right], \end{aligned} \tag{5}$$

where $(i, j) \in OD, k \in K(i, j), (l, m) \in S, \Delta t$ is the length of the time step in hours, $Le(l, m)$ is the length of the section in km, and q_{in} and q_{out} are respectively the flow into and the flow out of the section in the interval $(t, t + 1)$. These flows depend on the section considered, they are specified below for several special cases.

The transition of the average speed of section $(l, m) \in S$ is assumed to be given by the formula

$$\begin{aligned} v((l,m),t+1) \quad = \quad & v((l,m),t) + \frac{\Delta t}{\tau}[v^e(x((l,m),t)) - v((l,m),t)] \\ & + \frac{\Delta t}{L_e(l,m)} v((s,l),t)[v((s,l),t) - v((l,m),t)] \\ & + \Delta t(Le(l,m)lane(l,m))^2[\beta x((l,m),t) + (1-\beta)x((m,r),t)] \times \\ & \times [x((m,r),t) - x((l,m),t)] \end{aligned} \tag{6}$$

where $\tau \in (0, \infty)$ is a time constant, $\beta \in (0, 1)$, (s, l) is the adjacent upstream section, and (m, r) the adjacent downstream section. When either node l or node m is an intersection then the terms of (6) are set to zero that refer to an upstream or downstream section beyond the intersection. Here $v^e : R_+ \rightarrow R_+$ is the equilibrium relation between density and average speed. The actual form for this relation suggested by S.A. Smulders [16] will be used,

$$v^e(x) = \begin{cases} v_{free} - ax, & 0 \leq x \leq x_{crit}, \\ b[\frac{1}{x} - \frac{1}{x_{jam}}], & x_{crit} \leq x \leq x_{jam}, \\ 0, & x_{jam} \leq x. \end{cases} \tag{7}$$

For a particular stretch of freeway the parameter values of this function are $x_{crit} = 27$ veh/km.lane, $x_{jam} = 110$ veh/km.lane, $a = 0.58km^2/h$, $b = 3197veh/h.lane$. The second term on the right-hand side of (6) is called the *relaxation term.* If the average speed differs from the equilibrium speed according to the actual density via the equilibrium relation, then the average speed is adjusted. The third term is called the *convection term.* If in the upstream section the average speed is higher than in the section under consideration then the average speed is adjusted accordingly. The last term is called the *anticipation term.* If drivers notice that density is gradually increasing they are expected to anticipate on this by reducing their speed. This term is the most controversial one, several variants have been proposed. Additional validation of this term is needed.

Transition for a section not adjacent to an intersection Consider a section $(l, m) \in S$ such that at node $m \in I$ there is no intersection. Let $(i, j) \in OD$, $k \in K(i, j)$, and let $(m, s) \in S$ be the section downstream of (l, m). Then the flow out of section (l,m) for these variables is

$$\begin{aligned}
&q_{out}((i,j),k,(l,m),t) \\
&= \alpha \left[q((l,m),t) - \sum_{i_1 \in I_0, k_1 \in K(i_1,m)} p((i_1,m),k_1,(l,m),t) q((l,m),t) \right] \\
&\quad +(1-\alpha) \left[q((m,s),t) - \sum_{j_1 \in I_D, k_1 \in K(m,j_1)} \lambda((m,j_1),k_1,t) \right] \\
&\quad +p((i,m),k,(l,m),t) q((l,m),t) I_{\{j=m\}}. \qquad (8)
\end{aligned}$$

If $j \neq m$ then $q((l, m), t)$ represents the flow of section (l, m),

$$\sum_{i_1 \in I_0, k_1 \in K(i_1,m)} p((i_1,m),k_1,(l,m),t) q((l,m),t)$$

represents the exit flow at node m, and

$$\sum_{j_1 \in I_D, k_1 \in K(m,j_1)} \lambda((m,j_1),k_1,t)$$

represents the flow at node m into section (m, s) from outside the network. Here $\alpha \in [0, 1]$ is a weighting factor. The expression (8) is similar to that of (4). If $j = m$ then there is a flow out of the section for state $x((i, m), k, (l, m), t)$ as indicated in (8). The flow into section (m, s) is

$$\begin{aligned}
&q_{in}((i,j),k,(m,s),t) \\
&= \alpha \left[q((l,m),t) - \sum_{i_1 \in I_O, k_1 \in K(i_1,m)} p((i_1,m),k_1,(l,m),t) q((l,m),t) \right] \\
&\quad +(1-\alpha) \left[q((m,s),t) - \sum_{j_1 \in I_D, k_1 \in K(m,j_1)} \lambda((m,j_1),k_1,t) \right] \\
&\quad +\lambda((m,j),k,t) I_{\{i=m\}}. \qquad (9)
\end{aligned}$$

If there is no exit or no entry at node m then the corresponding terms in (9) are zero. It is assumed that for the OD-pair (i, j) there is an initial assignment of traffic to a route $r((i, j), k) \in R(i, j)$ with route number $k \in K(i, j)$. For example, route $r((i, j), k)$ may be selected such that the travel distance along route $r((i, j), k)$ is the smallest of all those of routes in $R(i, j)$. Below the model will be extended such that vehicles can change route.

Transition for a section adjacent to an intersection Consider section $(l, m) \in S$ and suppose that there is an intersection at node $m \in I$. The flow out of the section for $(i, j) \in OD$, $k \in K(i, j)$, and $(l, m) \in S$ is

$$q_{out}((i, j), k, (l, m), t) = p((i, j), k, (l, m), t)q((l, m), t)$$

where $(m, s) \in S$ is the section directly downstream from node m according to route k.

Consider next a section $(m, s) \in S$ that starts at a freeway intersection $m \in I_I$. For $(i, j) \in OD$, $k \in K(i, j)$ the flow into variable $x((i, j), k, (m, s), t)$ is given by

$$q_{in}((i, j), k, (m, s), t) = p((i, j), k, (l, m), t)q((l, m), t)$$

where (l, m) is the section directly upstream from (m, s) according to route k.

2.5 Measurements

The control and signalling system provides information on passage times and passages speeds of cars at measurement locations. This information is aggregated at the location in a detector station. Communicated from the detector station to the central computer are say 1 minute averages of the intensity and the average speed. This information may be used to estimate the traffic densities.

2.6 Boundary conditions

The intensities of the entry flows are assumed to be specified by the user of the model and not to depend on the state of the network. Consider next a section from which traffic leaves the network. The flow out of such a section $(l, j) \in S$ for $(i, j) \in OD$, $k \in K(i, j)$ is assumed to be given by

$$q_{out}((i, j), k, (l, j), t) = p((i, j), k, (l, j), t)q((l, j), t).$$

2.7 The full state space model

Consider the state variables $x((i, j), k, (l, m), .) : T \to R_+$, $v((l, m), .) : T \to R_+$, for $(i, j) \in OD$, $k \in K(i, j)$, and $(l, m) \in r((i, j), k)$. Collect these variables in the state vector $x : T \to R_+^M$ for some $M \in Z_+$. Similarly combine the arrival rates $\lambda((i, j), k, .) : T \to R_+$ into a vector $\lambda : T \to R_+^S$. These rates are assumed to be specified by the user of the model. The dynamic system may then be written as

$$x(t + 1) = f(x(t), \lambda(t)), x(t_0) = x_0. \tag{10}$$

The differences can now be pointed out between the discrete-time model of this paper and that of M. Papageorgiou and co-workers [8, 11]. In [11] a density per section is used. Then on page 477 the density is decomposed into a density per destination and a recursion is proposed for this density. However, it is not stated in the paper that this density per destination is used in the remainder of the paper or whether the heuristic approach of page 476 is used. In this heuristic approach one divides the density into fixed fractions. For the transition of these fractions a heuristic rule is proposed. In [8] the same model is proposed as in [11]. In the model of this paper state variables are considered consisting of densities per OD-pair, per route, and per section. As mentioned before, the dependence on the route prevents unrealistic rerouting. Experience with modelling will have to show the usefulness of the state space model and answer the question of minimality of the state space.

Example 2.2 *A state space model for a network with two routes.* Consider the freeway network of figure 2. The sets of node numbers are

$$I = \{1, 2, 3, 4, 5, 6\}, I_O = \{1\}, I_D = \{6\}, I_I = \{2, 5\}.$$

The set of OD-pairs is $OD = \{(1,6)\}$. The set of sections is

$$S = \{(1,2),(2,3),(3,5),(2,4),(4,5),(5,6)\}.$$

The set of routes for OD-pair (1,6) is

$$\begin{aligned}
r((1,6),1) &= \{(1,2),(2,3),(3,5),(5,6)\}, && (11)\\
r((1,6),2) &= \{(1,2),(2,4),(4,5),(5,6)\}, && (12)\\
R(1,6) &= \{r((1,6),1), r((1,6),2)\} = \{r((1,6),k), \quad k \in K(1,6)\}, && (13)\\
K(1,6) &= \{1,2\}. && (14)
\end{aligned}$$

The density state variables are

$$x((1,6),1,(1,2),t), x((1,6),1,(2,3),t), x((1,6),1,(3,5),t), x((1,6),1,(5,6),t), \quad (15)$$
$$x((1,6),2,(1,2),t), x((1,6),2,(2,4),t), x((1,6),2,(4,5),t), x((1,6),2,(5,6),t). \quad (16)$$

The state transitions are specified by the recursions

$$\begin{aligned}
x((1,6),1,(1,2),t+1) &= x((1,6),1,(1,2),t) + \frac{\Delta t}{Le(1,2)lane(1,2)} \times \\
&\quad \times[\lambda((1,6),1,t) - p((1,6),1,(1,2),t)q((1,2),t)], && (17)\\
q((1,2),t) &= lane(1,2)x((1,2),t)v((1,2),t), && (18)\\
x((1,2),t) &= x((1,6),1,(1,2),t) + x((1,6),2,(1,2),t) && (19)\\
x((1,6),2,(1,2),t+1) &= x((1,6),2,(1,2),t) + \frac{\Delta t}{Le(1,2)lane(1,2)} \times \\
&\quad \times[\lambda((1,6),2,t) - p((1,6),2,(1,2),t)q((1,2),t)], && (20)\\
x((1,6),1,(2,3),t+1) &= x((1,6),1,(2,3),t) + \frac{\Delta t}{Le(2,3)lane(2,3)} \times \\
&\quad \times[p((1,6),1,(1,2),t)q((1,2),t) \\
&\quad -\alpha q((2,3),t) - (1-\alpha)q((3,5),t)], && (21)\\
q((2,3),t) &= lane(2,3)x((1,6),1,(2,3),t)v((2,3),t), && (22)\\
x((1,6),1,(3,5),t+1) &= x((1,6),1,(3,5),t) + \frac{\Delta t}{Le(3,5)lane(3,5)} \times \\
&\quad \times[\alpha q((2,3),t) + (1-\alpha)q((3,5),t) - q((3,5),t)], && (23)\\
q((3,5),t) &= lane(3,5)x((1,6),1,(3,5),t)v((3,5),t), && (24)\\
x((1,6),2,(2,4),t+1) &= x((1,6),2,(2,4),t) + \frac{\Delta t}{Le(2,4)lane(2,4)} \times \\
&\quad \times[p((1,6),2,(1,2),t)q((1,2),t) \\
&\quad -\alpha q((2,4),t) - (1-\alpha)q((4,5),t)], && (25)\\
q((2,4),t) &= lane(2,4)x((1,6),2,(2,4),t)v((2,4),t), && (26)\\
x((1,6),2,(4,5),t+1) &= x((1,6),2,(4,5),t) + \frac{\Delta t}{Le(4,5)lane(4,5)} \times \\
&\quad \times[\alpha q((2,4),t) + (1-\alpha)q((4,5),t) - q((4,5),t)], && (27)\\
q((4,5),t) &= lane(4,5)x((1,6),2,(4,5),t)v((4,5),t), && (28)\\
x((1,6),1,(5,6),t+1) &= x((1,6),1,(5,6),t) + \frac{\Delta t}{Le(5,6)lane(5,6)} \times \\
&\quad \times[q((3,5),t) - p((1,6),1,(5,6),t)q((5,6),t)], && (29)\\
q((5,6),t) &= lane(5,6)x((5,6),t)v((5,6),t), && (30)
\end{aligned}$$

$$x((5,6),t) = x((1,6),1,(5,6),t) + x((1,6),2,(5,6),t) \quad (31)$$

$$x((1,6),2,(5,6),t+1) = x((1,6),2,(5,6),t) + \frac{\Delta t}{Le(5,6)lane(5,6)} \times \times[q((4,5),t) - p((1,6),2,(5,6),t)q((5,6),t)]. \quad (32)$$

3 Filtering of freeway traffic flow

The solution to the routing problem to be considered in the next section requires information on the actual state of traffic in the network. Therefore the filtering problem for the dynamic system of the previous section is of interest. The filtering problem asks for an algorithm to estimate the state of traffic in the network based on measurements of passage times and passage speeds of cars at detector stations. A filter for this purpose of a corresponding continuous-time system has been developed by S.A. Smulders, see [17, 18]. The filter is based on an extended Kalman filter with counting process measurements. Suggestions have been made to simplify the filter of [17] to reduce its complexity. The filtering problem will not be discussed any further in this paper.

A second problem to be discussed in this section is the estimation of traffic intensities for the OD-flows given data on the flows through the links, entries, and exits. These intensities are needed in the model and are not directly measured. There are techniques to estimate the entries in this matrix. Since traffic intensities are in practice not constant but time varying, techniques have been proposed to estimate slowly varying arrival intensities of origin-destination flows [4]. Adaptive estimation of these intensities may be useful for the solution of the routing problem.

4 Routing

The routing problem has been motivated in section 1. The general aim of routing is to minimize travel cost. In this section the routing problem is formulated and two subproblems are analysed.

Consider a freeway network on which a control and signalling system is installed. Suppose that it is possible to provide drivers with route guidance. For example, at every freeway intersection it is possible to display variable directions signs for each destination. Or it may be possible to inform drivers by radio or by another communication channel that transmits messages from the control center to drivers. In the network of example 2.1 drivers on the link (1,5) for the OD-pair (1,3) may at intersection 5 be advised to follow route $r((1,3),1)$ or route $r((1,3),2)$, see (1,2).

4.1 How variable direction signs affect the traffic flow in the network

The state space model for freeway traffic flow in a network as proposed in section 2 may be considered as a control system. To complete the specification of this system the effect of the input on the flows must be specified. This is first done for an example.

Consider example 2.2 with figure 2. For the OD-pair $(1,6) \in OD$ there are two routes, r((1,6),1) and r((1,6),2). These routes have the sections (1,2) and (5,6) in common. There is a separation of the routes at the intersection of node 2. Therefore a routing advice may be displayed at intersection 2. The effect of this routing advice on the flow may be modelled in the state transitions by

$$x((1,6),1,(2,3),t+1) = x((1,6),1,(2,3),t) + \frac{\Delta t}{Le(2,3)lane(2,3)}[u((1,6),1,2,t)q((1,2),t) - \alpha q((2,3),t) - (1-\alpha)q((3,5),t)], \quad (33)$$

$$x((1,6),2,(2,4),t+1) = x((1,6),2,(2,4),t) + \frac{\Delta t}{Le(2,4)lane(2,4)}[u((1,6),2,2,t)q((1,2),t) - \alpha q((2,4),t) - (1-\alpha)q((4,5),t)]. \quad (34)$$

Here $u((1,6),1,2,.) : T \to [0,1]$ represents the fraction of the flow for OD-pair (1,6) that at intersection 2 is directed to route 1 and $u((1,6),2,2,.) : T \to [0,1]$ the fraction directed to route 2. The relation $u((1,6),1,2,t) + u((1,6),2,2,t) = 1$ should hold for all $t \in T$. If $u((1,6),1,2,t) = 1$ or 0 then traffic is directed at route 1 or route 2 respectively. It will be argued below that it is useful to consider also intermediate values for u, or to let $u((1,6),1,2,t) \in [0,1]$.

In general the effect of routing control on the flow in a freeway network is described as follows. Consider the OD-pair (i,j). The set of routes for this OD-pair is $R(i,j) = \{r((i,j),k), k \in K(i,j)\}$. At each intersection where the traffic flow for (i,j) branches into two or more routes a routing advice may be displayed. If the intersection at node m is such a branching point then the freeway operator must specify the routing fractions $u((i,j),k,m,.) : T \to [0,1]$. These fractions must satisfy the condition that for all $(i,j) \in OD$, $m \in I$, $t \in T$ fixed

$$\sum u((i,j),k,m,t) = 1,$$

where the sum is over all routes that branch at intersection m. In addition the routing fractions have to be consistent, all traffic destined for node j has to be advised to follow the same route. In the state transition equations the effect of routing control is modelled as

$$u((i,j),k,m,t)[\sum_{k_1} p((i,j),k_1,(l,m),t)q((l,m),t)]$$

where the sum is over all routes numbered k_1 that use section (l,m) followed by section (m,s_1), in which node s_1 depends on route k_1.

Let $u : T \to U$ be the vector of all routing fractions for all $(i,j) \in OD$ and all relevant freeway intersections. The input $u(t)$ describes the set of all routing fractions; it will be called a *route scheme* or a *route input*. The routing control system may then be written as

$$x(t+1) = f_1(x(t),u(t),\lambda(t)), \quad x(t_0) = x_0. \tag{35}$$

A *routing law* is a map $g : X \to U$ that for each state x specifies the routing fractions in the network. A routing law specifies the input to the control system via

$$u(t) = g(x(t)). \tag{36}$$

The control system (35) with the control law (36) then becomes a routing system given by

$$x(t+1) = f_1(x(t),g(x(t)),\lambda(t)) = f_2(x(t),\lambda(t)), \quad x(t_0) = x_0. \tag{37}$$

It is suggested that in route guidance non-integer values of the routing fractions are useful. Consider first the case of integer routing fractions, in particular the case in which for an OD-pair there exist only two routes. At the branching point the input space is $U = \{0,1\}$, and the input is $u((i,j),k,m,t) = 1$ or 0, corresponding to the routing advice that directs traffic to route 1 or to the alternate route. During periods of congestion when the future travel time along both routes is approximately equal, the input is likely to oscillate between both routes. In fact, the simulations mentioned in the paper [11] show precisely this fact. Such oscillations seem detrimental to the stability of the traffic flow. In communication networks this phenomenon is less of a problem due to the faster dynamics.

With an input that may take non-integer values traffic is not likely to exhibit such oscillations. If the input space and the input at an intersection are $U = [0,1]$ and $u((i,j),k,m,t) = 0.7$, say, then road users are advised to separate out such that 70% follows route k and 30% follows the alternate route. Experiments could be performed whether road users are willing to follow such an advice. In periods of congestion the input will no longer be wildly oscillating. In practice one may want to limit the number of input values, say to $U = \{0,0.3,0.5,0.7,1.0\}$.

Following [11] a compliance rate may be introduced. This rate is defined as the fraction of car drivers that will follow the displayed routing direction. The compliance rate will be route and location dependent. Values of the compliance rates are currently unknown. If routing is introduced as a regular means of control on freeways then the agency in charge of this operation will have to educate the drivers and to convince them of the advantage of routing. Since the investigation on which this paper reports is a feasibility study, the compliance rates are set to one. This means that all drivers are supposed to follow the route guidance.

4.2 The routing problem

The control objectives of routing are:

1. Minimize the travel cost of network users.
2. Provide fair service to all network users.

In general one would want to minimize travel costs, consisting of travel time, energy used, etc. Of these costs the travel time is considered as dominant, hence travel time will be used in the remainder of this paper.

Because the routing problem is like a multidecisionmaker problem one should distinguish between the decision criteria of *user optimality* and of *network optimality*. These concepts were introduced in [22] with the names of *user optimum* and *network optimum*. In user optimality for a given OD-pair, a route is selected for this OD-pair on the basis of the first control objective while keeping the routes of the other OD-pairs fixed. Although in principle the road users make the choice of a route, the choice will be the same for each user travelling on the OD-pair assuming that they are provided with the same information and assuming rational behaviour. In network optimality the choice of routes for all OD-pairs is made on the basis of both control objectives for the complete network.

The synthesis of a routing law will in the first stage of the investigation be based on the separation principle. According to this principle a routing law may be synthesized by combining a filter algorithm with a routing law based on state feedback. Therefore attention will first be concentrated on synthesis of a routing law under the assumption that the state of the control system is observed. The synthesis of a routing law asks questions such as: What part of the state vector should be used in the routing law? What is the analytic form of the routing law? How does the analytic form of the routing law influence the performance of the control system? The approach to answering these questions will be to propose routing laws based on engineering experience and thinking, to evaluate proposed routing laws, and to use optimal control theory to derive structural properties of routing laws.

Below the routing problem will be considered for both user optimality and network optimality.

4.3 Routing for the decision criterion of user optimality

Consider a freeway network and an OD-pair. The control objective is to minimize the travel cost for this OD-pair. As indicated above, in a first approach attention is restricted to control based on state feedback.

Problem 4.1 *The* optimal routing problem with as decision criterion user optimality *is to select for each OD-pair a routing law that will minimize the travel cost while keeping the route scheme for the other OD-pairs fixed.*

The main question for problem 4.1 is on what part of the state vector the routing law depends.

Example 4.2 Consider the two route network displayed in figure 2. For the OD-pair (1,6) the set of routes is $R(1,6) = \{r((1,6),1), r((1,6),2)\}$. At intersection 2 routing advice may be displayed to follow route r((1,6),1) via node 3 or route r((1,6),2) via node 4. The input variable is the branching at node 2. The routing law that determines these branching fractions may depend on the state of the network.

Example 4.3 Consider the freeway network of figure 1 and the traffic flow for OD-pair (1,3). The set of routes for this OD-pair is $R(1,3) = \{r((1,3),1), \quad r((1,3),2)\}$. At intersection 5 the routes branch, hence a routing advice may be displayed there. It will be assumed that for other OD-pairs the routes are kept constant during the analysis.

Overview of literature on the subject

The problem of this subsection is the routing problem with the decision criterion of user optimality. The classical approach of route selection given arrival intensities by optimization techniques does not make use of the actual state information of the network and is therefore not suitable. Attention is restricted to feedback control laws that use actual state information. Several synthesis approaches have been suggested, see [2, 3, 9, 11].

M. Papageorgiou [11] has proposed to use optimal control theory to synthesize a control law that meets the control objectives. See that paper for an example and for a simulation of the controlled system. The conclusion of that paper is that the routing advice is highly oscillating between the alternate routes. Other disadvantages are the high computational cost of the optimal control law and that the structure of the control law is hard to determine. If the arrival intensities of the network change then the control law has to be recomputed. Therefore this approach has its limitations. Papageorgiou in [11] has also proposed linearization of the nonlinear control law and application of a time-invariant LQG-control law. This approach has serious disadvantages because in the state of traffic congestion the control system is highly nonlinear. Papageorgiou has also proposed a state feedback law consisting of a proportional and integral term. The parameters of this control law must be determined by trial and error. In the paper [11] a simulation is shown of the control system with this control law. Difficulties with this approach are the computation of the control law parameters and the suitability of a linear control law for a nonlinear control system. Therefore this synthesis approach has also its limitations. Nevertheless, the author values the pioneering work of Papageorgiou as presented in [11].

The analogy of routing in a freeway network with routing in a communication network may be exploited. Consider an elementary model of routing in a communication network consisting of an arrival stream and two service stations. Tasks arriving at the routing switch must be assigned to one of two service stations. If a task is assigned to a service station then it will be served directly or put in the queue and served at a later time. Let x_1, x_2 represent the queue lengths at the stations 1 and 2. The control objective is to minimize the waiting and service times of all tasks. Under certain assumptions it may be shown that the optimal control law for this control objective is based on a partitioning of the state space $X = N \times N$ into two connected components by a curve, say $x_2 = h(x_1)$. For example, the control is such that if on arrival of a task the state of the queueing system is such that $x_2(t) \leq h(x_1(t))$ then the task is assigned to station 2 else to station 1. This control law is nonlinear. For references on this approach see [5, 21]. In the result quoted above the queue lengths represent the waiting times at both stations. The implications of this result for routing in a freeway network is that that route is selected that yields the lowest future travel time. This conclusion is of course dependent on the cost criterion considered.

Proposed approach to routing for user optimality

Restrict attention then to the case in which the travel cost is equal to the travel time. Then the routing law will direct traffic to the route with the lowest future travel time. A major question is then to estimate the future travel time for all admissable routes. Note that this estimate will in general depend on the future input values, also for routes selected for other OD-pairs. By which procedure may one estimate the future travel time along a route? No general procedure is known.

A simple estimate of the future travel time along a route may be obtained by simulation. Consider a chain of sections

$$\{(i_1, i_2), (i_2, i_3), \ldots, (i_{N-1}, i_N)\},$$

and the traffic flow system

$$x(t+1) = f(x(t), \lambda(t)), \quad x(t_0) = x_0.$$

Suppose a car enters the first section at time $\tau_0 \in T$. The time at which it will on average exit from the first section is

$$\tau_1 = \tau_0 + \frac{Le(i_1, i_2)}{v^e(x((i_1, i_2), \tau_0))},$$

where the second term on the right-hand side is the length of the section divided by the equilibrium speed associated with the density in the section at time τ_0. This expression is an approximation because the average speed may change in the interval $[\tau_0, \tau_1]$. Because $(\tau_1/\Delta t) \in R_+$ may not be an integer let $[\tau_1/\Delta t]$ be the integer nearest to $\tau_1/\Delta t$. An estimate of the travel time of the car in the second section is then

$$\frac{Le(i_2, i_3)}{v^e(x((i_2, i_3), [\tau_1/\Delta t]))}.$$

Proceeding this way an estimate of the travel time along route k is

$$\hat{\tau}(k) = \sum_{n=1}^{N-1} \frac{Le(i_n, i_{n+1})}{v^e(x((i_n, i_{n+1}), [\tau_{n-1}/\Delta t]))}.$$

This estimate may be computed by simulation of the freeway network.

A second estimate for the future travel time along a route may be based on the current state of traffic along the route. Consider a branching point of two routes for an OD-pair. Denote the route from the branching point onwards by

$$r((i,j), k) = \{(i_1, i_2), (i_2, i_3), \ldots, (i_{N-1}, i_N)\}.$$

An estimate of the future travel time along route k if the route input directs traffic to route k is

$$\hat{\tau}(k) = \sum_{n=1}^{N-1} \frac{Le(i_n, i_{n+1})}{v^e(x((i_n, i_{n+1}), t))}. \tag{38}$$

This estimate uses only the current state of traffic and neglects the dynamics of the traffic in the network. If $K(i,j) = \{1, 2\}$ then the routing law should be such that if $\hat{\tau}(1) < \hat{\tau}(2)$ then route 1 is selected else route 2. This routing law may be written as

$$u((i,j), 1, m, t) = \begin{cases} 1, & \text{if } \hat{\tau}(1) < \hat{\tau}(2), \\ 0, & \text{otherwise}, \end{cases}$$

$$u((i,j), 2, m, t) = \begin{cases} 0, & \text{if } \hat{\tau}(1) < \hat{\tau}(2), \\ 1, & \text{otherwise}, \end{cases}$$

and as $u(t) = g(x(t))$. This is a nonlinear control law.

An even simpler estimate of the future travel time may be based on only the length of the route and the length of the traffic jams along the evaluated route. Let the length of the route be L_1 km and the length of the traffic jams along the route be L_2 km. An estimate of the travel time is then

$$\hat{\tau} = \frac{L_1 - L_2}{v_1} + \frac{L_2}{v_2},$$

where v_1 is the speed if there is no traffic jam, say $v_1 = 90\ km/h$, and v_2 is the speed with a car passes a traffic jam, say $v_2 = 10\ km/h$. This estimate is rather crude but considering the dynamics of traffic flow in a network it may not be unrealistic. For example, if $L_1 = 18$ km, $L_2 = 3$ km, $v_1 = 90$ km/h, and $v_2 = 10$ km/h then

$$\hat{\tau}_1 = \frac{18-3}{90} + \frac{3}{10} = 28 \quad \text{minutes} \; ,$$

while if there is no traffic jam then $\hat{\tau}_2 = 18/90 = 12$ minutes.

A mathematical approach to the estimation of future travel time may be formulated by analogy with queueing networks. A stretch of freeway may be modelled as a chain of server stations. Each station processes tasks according to the first-in-first-out discipline. Upon arrival of a task at the first station the length of the queues at all stations is observed. One may compute the conditional expectation of the future travel time given the current state of the network. The result depends on the assumption made for the server distribution. One may also question the suitability of the queueing network model for a freeway stretch. This approach will not be detailed in this paper.

4.4 Routing with as decision criterion network optimality

Consider the routing problem with as decision criterion network optimality. The control objectives of routing are to minimize the total travel cost and to provide a fair service to all users of the network. One way to put these criteria in one expression is to consider the weighted future travel time, in which the weights are proportional to the traffic intensities of the respective OD-pairs. As stated above, attention is first restricted to control based on state feedback. With the travel cost or travel time one may associate a cost function.

Problem 4.4 *The* routing problem with as decision criterion network optimality *and with the class of state feedback routing laws is to select that routing law that minimizes the travel cost and that provides fair service to all network users.*

The synthesis of routing laws may be approached by optimal control, or by engineering reasoning, or by a combination of these. For the optimal control approach it is in principle possible to write the dynamic programming equation for the value function. The dynamic programming equation for the value function will include, for any state, a minimization over all possible routes. There are methods and techniques to approximate the value function numerically by discretization of the state space. It should be clear that even for relatively simple networks the complexity of the computations is very high. It is therefore of interest to limit the class of routing laws, and hence the complexity of the routing problem. It does not seem realistic to synthesize a routing law for all states.

For the network of figure 1 there are 8 origin and destination nodes. The origin 1 has for each of the destinations $\{2, 3, 4, 9, 10, 11, 12\}$ two possible routes, although some of the alternate routes, like for $(1, 9) \in OD$, are unrealistic. In total there are thus 2^7 route schemes. For the four origins $\{1, 2, 3, 4\}$ there are 4.2^7 route schemes. In a dynamic programming equation one would have to minimize over a set with as many elements.

For the synthesis of a routing law the following synthesis procedure of nonlinear control is proposed. Suppose that the arrival intensities are constant during the period considered, $\lambda(t) = \lambda_1$. The control system is then given by

$$x(t+1) = f(x(t), u(t), \lambda_1), \;\; x(t_0) = x_0.$$

If the arrival intensities are relatively low then traffic is directed to those routes that have the shortest travel distance. Denote this route scheme by $u_1 \in U$. Then the equilibrium state $x_e \in X$ is given by the equation

$$x_e = f(x_e, u_1, \lambda_1).$$

For arrival intensities that are relatively high, the route scheme for the equilibrium state may be determined by an optimization technique.

Actual freeway traffic flow is rather irregular. This irregularity is due to fluctuations in the arrivals and to the variety of driver behaviour. Therefore the state will fluctuate around the equilibrium state and a stability analysis is necessary.

Next a stability analysis must be made of the behaviour of the traffic system at the equilibrium state. It is conjectured that if the density in every section is below the critical density of 27 veh/km.lane, that then the equilibrium state is stable. If in any section the density is higher than the critical value mentioned then another routing scheme should be considered.

How to route traffic if a traffic jam has formed in the network? The complexity of the state space model directs attention to a finite valued control law of the form

$$g(x) = \sum_{i=1}^{N} u_i I_{(x \in X_i)}, \tag{39}$$

where $\{X_i, i = 1, \ldots, N\}$ is a partition of the state space and $u_1, \ldots, u_n \in U$. Thus for each part of the state space a constant input is used. If the routing fractions are only allowed to take the values 0,1 then the routing law will naturally have the form displayed above. The routing law proposed is to have a similar form but the number N of possible routing laws should be much smaller than if the routing fractions take only the values 0 or 1. For example, if a traffic jam has formed in a link then one tries to reroute traffic such that the jam may be dissolved and the state of the network may return to the equilibrium state. A control law of the form (39) is used in bang-bang control in which the input space is say $U = [-1, +1]$.

Synthesis of a control law of the form (39) demands a partition of the state space and an input for each part of the partition. As to the partition, no concrete suggestions will be made. However, it is suggested to periodically compute a new input given the current state of traffic. This input can then be held constant for a fixed time or till the state of traffic has entered a neighborhood of the equilibrium state. The main question therefore is, given the state of traffic in the network, how to select a routing scheme such that the state will return to the equilibrium state? The approach to this problem is to modify the nominal route scheme, the route scheme associated with the equilibrium state. This approach is detailed below for a particular state of traffic.

Restrict attention then to a particular state of traffic. Suppose that congestion has developed in one link of the network and that other links are free of congestion. Consider the freeway network of figure 1 and let the congested link be (10,7). The following algorithm for the selection of a route scheme is proposed for this case.

Algorithm 4.5 Route selection in case of one congested link for the network of figure 1.
Suppose given the information on the freeway network and on the state of traffic in the network.

1. *Determine all currently used routes that include the congested link. Determine also the associated OD-pairs.*

2. *Of the OD-pairs determined in step 1, determine those that have an alternative route not including the congested link. Determine also the associated route.*

3. *Order the routes determined in step 2 increasingly according to the relative length of the alternate route d_2/d_1, where d_1 represents the length of the nominal route and d_2 represents the length of the alternate route.*

4. *For each of the following route schemes estimate the average future travel time:*

    - *For the currently used nominal route.*
    - *For the route scheme obtained from the nominal route scheme by changing it increasingly by the routes determined in step 3.*

5. *Select that route scheme of step 4 that minimizes the weighted average future travel time.*

Some remarks on the above algorithm are in order. The dynamic programming algorithm requires for any state a minimization over all route schemes. The above algorithm limits the search from all routes to a smaller subset of routes. The algorithm neglects that rerouting flows away from the congested link may require rerouting flows not currently using the congested link. The purpose of rerouting flows currently using the congested link is that by a reduction of the flow to this link the congestion will dissolve and that then traffic may be brought back to normal operating conditions. The seemingly naive approach of ordering in step 3 is motivated by the argument that in the comparison of routes the estimated future travel time of the alternate route is likely to be shorter if the relative length of the route is smaller. Recall that attention is restricted to the case in which only one link is congested.

Example 4.6 Consider again the freeway network of figure 1. Suppose that there is congestion in link (10,7). The algorithm 4.5 will be applied to select a new route scheme.

1. Routes that are using the congested link (10,7) could be those for the OD-pairs (1,3), (2,3), (2,4), and (2,11). For example, for (1,3) the route could be

   $$r((1,3),1) = \{(1,5),(5,9),(9,6),(6,10),(10,7),(7,3)\}.$$

   If variable direction signs were also placed at the freeway entrances at the nodes 9 and 10 then routes for four other OD-pairs would have to be counted: (9,3), (9,11), (10,4), and (10,12).

2. All the OD-pairs mentioned in step 1 have an alternative route not including link (10,7).

3. According to the assumption that each link has length 1 the relative lengths are: (1,3) 1.0, (2,4) 1.0, (2,11) 1.5, and (2,3) 2.0. In general it does not seem realistic to advise rerouting if the relative length exceeds the value 1.5 or the value 2.0.

4. Estimate the future travel time for the route schemes:
   - The nominal route.
   - The route obtained from the nominal route by replacing $r((1,3),1)$ by $r((1,3),2)$.
   - The route obtained from that of the previous step by replacing the route for $(2,4) \in OD$ that includes link (10,7) by the alternate route not including this link.
   - As in the previous step but for the $(2,11) \in OD$.

5. Select the route scheme of step 4 that minimizes the weighted future travel time.

It remains to evaluate how useful this algorithm is. The synthesis question formulated above may also be approached by the study of simple networks.

5 Conclusions and open problems

The contributions of this paper to routing of traffic in a freeway network are the state space model and the approaches to the routing problem. A state space model for traffic flow in a freeway network has been proposed. The state variables are distinguished per section and consist of the density per origin-destination pair and per route, and the average speed of all cars in the section. The effect of variable direction signs on the traffic flow has been modelled.

The routing problem has been considered for the decision criteria of user optimality and of network optimality. Approaches to these problems have been discussed and several questions formulated.

Open problems

The following research problems seem of interest to the routing problem of freeway traffic flow.

1. Evaluation of the state space model proposed in this paper. The state transition relation for the average speed has to be evaluated for conditions near and in congestion.
2. The estimation of intensities for origin-destination flows from data about the flows through the links. An approach is indicated in [4]. Modelling of the time dependence of arrival intensities expecially for peak hour traffic.
3. The dynamic behaviour of a congested link. How does the head and the tail of a traffic jam move along the freeway for a range of arrival intensities? By how much must the arrival intensity be reduced and for how long, to clear a traffic jam?
4. Filtering of traffic data to obtain estimates of density and average speed per section. The approach of [17] may be used.
5. The estimation of future travel time. How to estimate the future travel time along a route? Collection and analysis of actual travel times. Calculation of the estimates proposed in section 4. Comparison of the estimates with values obtained by simulation and by measurement of actual travel time.
6. Synthesis of a routing law for the decision criterion of network optimality. Analysis of the nonlinear control system and of its dynamic behaviour.

References

[1] D. Bertsekas and R. Gallager. *Data networks.* Prentice-Hall Inc., Englewood Cliffs, 1987.

[2] C. Charbonnier, J.L. Farges, and J.J. Henry. Le routage dynamique des vehicules. In C. Commault et al., editor, *Proceeding First European Control Conference*, pages 2391–2396, Paris, 1991. Hermès.

[3] C. Charbonnier, J.L. Farges, and J.J. Henry. Models and strategies for dynamic route guidance. In Commission EC, editor, *Advanced Telematics in Road Transport - Proceedings DRIVE Conference - Brussels, February 4-6, 1991*, pages 106–112, Amsterdam, 1991. Elsevier.

[4] M. Cremer and H. Keller. A new class of dynamic methods for the identification of origin-destination flows. *Transpn. Res.-B*, 21:117–132, 1987.

[5] B. Hajek. Optimal control of two interacting service stations. *IEEE Trans. Automatic Control*, 29:491–499, 1984.

[6] U. Karaaslan, P. Varaiya, and J. Walrand. Freeway traffic flow, platooning and control. Report, Department of Electrical Engineering and Computer Science, University of California, Berkeley, 1990.

[7] M.J. Lighthill and G.B. Whitham. On kinematic waves ii. a theory of traffic flow on long crowded roads. *Proc. Roy. Soc. Series A*, 229:317–345, 1955.

[8] A. Messmer and M. Papageorgiou. METANET: A macroscopic simulation program for motorway networks. *Traffic Engineering + Control*, x:466–470, 1990.

[9] J.M. Morin, P. Gower, M. Papageorgiou, and A. Messmer. Motorway networks, modelling and control. In Commission EC, editor, *Advanced Telematics in Road Transport - Proceedings DRIVE Conference - Brussels, February 4-6, 1991*, pages 148–171, Amsterdam, 1991. Elsevier.

[10] M. Papageorgiou. *Application of automatic control concepts to traffic flow modeling and control.* Springer-Verlag, New York, 1983.

[11] M. Papageorgiou. Dynamic modeling, assignment, and route guidance in traffic networks. *Transpn. Res.-B*, 24B:471–495, 1990.

[12] H.J. Payne. Models of freeway traffic and control. *Simulation council*, 1:51–61, 1971.

[13] H.J. Payne. A critical review of a macroscopic freeway model. In *Engineering Foundation Conference on Research Directions in Computer Control of Urban Traffic Systems*, pages 251–265, 1979.

[14] S.A. Smulders. Modelling and simulation of freeway traffic flow. Report OS-R8615, Centre for Mathematics and Computer Science, Amsterdam, 1986.

[15] S.A. Smulders. Modelling and filtering of freeway traffic flow. In N.H. Gartner and N.H.M. Wilson, editors, *Proceedings of the 10th International Symposium on Transportation and Traffic Theory*, pages 139–158, New York, 1987. Elsevier.

[16] S.A. Smulders. Control of freeway traffic flow. Report OS-R8817, Centrum voor Wiskunde en Informatica, Amsterdam, 1988.

[17] S.A. Smulders. Filtering of freeway traffic flow. Report OS-R8806, Centrum voor Wiskunde en Informatica, Amsterdam, 1988.

[18] S.A. Smulders. Control of freeway traffic flow. Thesis, University of Twente, Enschede, 1989.

[19] S.A. Smulders. Control of freeway traffic flow by variable speed signs. *Transpn. Res.-B*, 24B:111–132, 1990.

[20] J.H. van Schuppen. Routing of freeway traffic - a state space model and routing problems. Preprint, CWI, Amsterdam, 1991.

[21] J. Walrand. *An introduction to queueing networks.* Prentice Hall, Englewood Cliffs, NJ, 1988.

[22] J.G. Wardrop. Some theoretical aspects of road traffic research. *Proceedings/Institution of Civil Engineers*, 1:325–378, 1952.

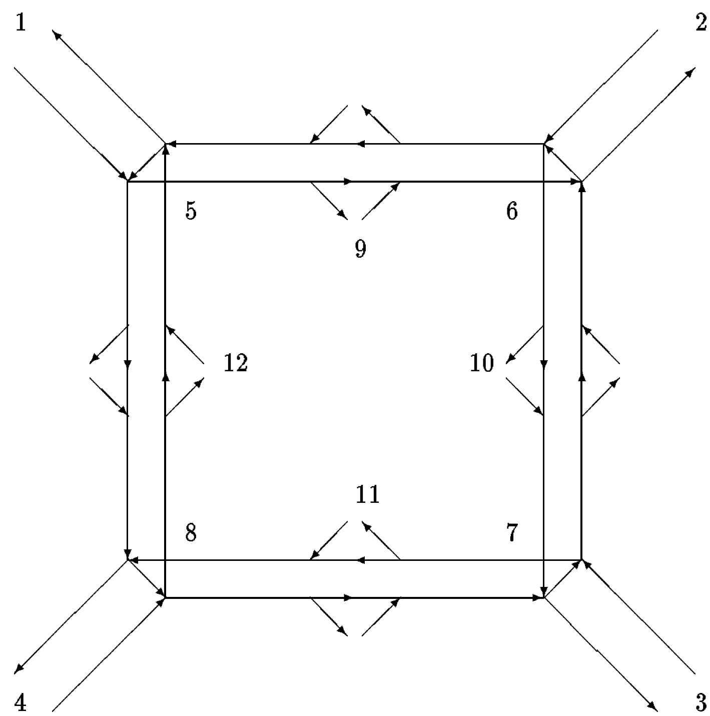

Figure 1: A freeway network in the form of a ring around an urban area.

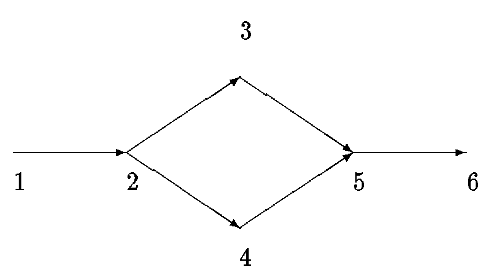

Figure 2: A freeway network consisting of two routes.