Control of day-to-day route choice in traffic networks with overlapping routes

M. van den Berg, B. De Schutter, H. Hellendoorn, and A. Hegyi

If you want to cite this report, please use the following reference instead:
Control of day-to-day route choice in traffic networks with overlapping routes

Monique van den Berg, Bart De Schutter, Hans Hellendoorn *
Andreas Hegyi **

* Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands, (e-mail: {monique.vandenberg,j.hellendoorn}@tudelft.nl,b@deschutter.info)
** Department of Transport & Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O.Box 5048, 2600 GA Delft, The Netherlands, (email: a.hegyi@tudelft.nl)

Abstract: We develop a day-to-day route choice control method that is based on model predictive control (MPC). For the route choice we assume that drivers base their decision on the experienced travel times. These travel times can be influenced via existing control measures, e.g. outflow limits or variable speed limits. This allows us to indirectly influence the route choice of the drivers. In previous work we have developed a route choice control method for networks with simple, non-overlapping routes, single destinations, and constant or piecewise constant flows. In this paper we extend this method to networks with overlapping routes and with restricted link inflow capacities. We illustrate the control approach with a case study for a simple network with four routes.

1. INTRODUCTION

Traffic demand is increasing rapidly, causing increased flows on nearly all roads. On roads that are in residential areas, near primary schools or shopping centers, a large flow can cause undesired or unsafe situations. Large flows also cause pollution and noise, which does not only affect humans living near the roads, but can also have a negative impact on nature reserves. Road administrators can influence the location of the large flows by changing the routes that drivers select by influencing the travel time on different routes, as proposed by Haj-Salem and Papageorgiou (1995); Taale and van Zuylen (1999). The travel times can be influenced indirectly using traffic control measures, e.g. traffic signals, ramp metering installations, or dynamic speed limits displayed on variable message signs.

We propose a control method that uses existing traffic control measures to influence the route choice of the drivers in a network where different routes are overlapping. We assume that drivers adapt their route choice for the next day based on the experienced travel time on the current day, as described by Bogers et al. (2005). As control method we use model predictive control (MPC) (Maciejowski, 2002). This is a model-based control approach that uses a prediction model in combination with an optimization algorithm to determine optimal settings for the traffic control measures. The optimal settings are then applied using a rolling horizon approach.

MPC has already been used previously to influence the route choice within a day (see e.g. Hegyi (2004); Bellemans (2003); Kotsialos et al. (2002); Peeta and Mahmassani (1995)). However, in this paper we focus on the day-to-day route choice. Day-to-day route choice is described by e.g. Maher (1998); Cascetta et al. (2000). The models used in those papers are complex, and require large computation times. As a result, these models are unsuitable for the use in model-based controllers. In previous work (van den Berg et al. (2008b,a)) we have formulated a basic route choice model for networks with non-overlapping routes and single destinations, and we have illustrated the possibilities of route choice control for these routes. In van den Berg et al. (2009) this model has been extended to include time-varying piecewise-constant demand patterns.

In the current paper we further extend the work of previous papers as follows: we complete the model by including overlapping routes and link capacities, and we develop a control method using the complete model. Note that the main purpose of the model developed in this paper is to be able to quickly assess the effect of control measures on route choice. For more detailed assessments more complex — and thus also more time-consuming — models should be used.

This paper is organized as follows. The control problem is formulated in Section 2. Next, the route choice model is formulated in Section 3, in particular the extensions for road capacities and overlapping routes. Section 4 presents the model-based control approach. In Section 5 the proposed control approach is illustrated with an example. Section 6 concludes the paper.

2. CONTROL PROBLEM

2.1 Network set-up

Throughout the paper we illustrate our approach based on the network shown in Fig. 1. The network has 1 origin and 1 destination, 1 intermediate vertex, 4 links, and 4 possible routes, as shown in Fig. 2.

Each link \( l \) (\( l \in L \)) in the network can be described by the following parameters, where \( d \) is the counter for the days. The length of link \( l \) is denoted by \( \ell_l \) (km), and its inflow capacity is denoted by \( Q_l^{in} \) (veh/h). The speed limit \( v_1(d) \) (km/h) gives the maximum speed that is allowed on link \( l \) at day \( d \). The outflow
Distributed over the routes according to the turning rates constant function during \( Q \). We assume that the demand are allowed to leave the link.

This means that each route can be characterized by its average value \( \beta \) between \( v \) and \( l_{\text{min}} \). The variable speed limits \( l \) use \( v \) instead of \( w \). Minimizing the cost function can have negative side effects. For example, reducing the flow on one route could in general lead to an increased flow on other routes, with congestion and longer travel times as results. This can be solved by introducing constraints on e.g. the travel times or flows. In general, typical constraints related to route choice are upper and/or lower bounds on the control variables, the flows, the travel times, and the queue lengths.

2.4 Constraints

Minimizing the cost function can have negative side effects. For example, reducing the flow on one route could in general lead to an increased flow on other routes, with congestion and longer travel times as results. This can be solved by introducing constraints on e.g. the travel times or flows. In general, typical constraints related to route choice are upper and/or lower bounds on the control variables, the flows, the travel times, and the queue lengths.

2.5 Overall control problem

The overall control problem can then be formulated as follows. Find the optimal control input that minimizes the cost function given the network, a route choice model, and the constraints.

3. ROUTE CHOICE MODEL

In this section we will extend the model of van den Berg et al. (2009) to include overlapping routes and link capacities.

3.1 Travel times

We assume that there are vertical queues in the network. This means that the drivers first drive through the link experiencing the free flow travel time. Next, they enter the vertical queue that is possibly formed due to the limited outflow capacity of the link. In this queue, if present, the drivers wait until they can leave the link, which leads to an experienced waiting time in the queue. This means that the average experienced travel time \( \tau_{\text{route}}(d) \) on a route \( r \) has two components: the free flow travel time on the route, and the total time spent in the queues:

\[
\tau_{\text{route}}(d) = \sum_{l \in L_r} \left( \tau_{\text{free}}(d) + \tau_{\text{queue}}(l) \right)
\]

where \( L_r \) is the set of all links \( l \) in route \( r \). The free flow travel time \( \tau_{\text{free}}(d) \) on link \( l \) is given by

\[
\tau_{\text{free}}(d) = \frac{L_l}{v_l(d)}.
\]

The computation of the average time spent in the vertical queue is more involved. We assume that the queue on a link is divided cost functions serve either to handle as much traffic as possible in a short time, or to keep vehicles away from protected routes (e.g. routes through residential areas or nature reserves). We will give two examples of possible cost functions. The weighted total travel time can be computed as follows

\[
j_{\text{TT}} = \sum_{d=1}^{N} \sum_{r \in R} \omega_r \beta_r(d) \left( \sum_{i=0}^{n} D(d,i) (t_{i+1} - t_i) \right) \tau_{\text{route}}(d)
\]

with \( R \) the set of all routes in the network, \( N \) the number of considered days, and weights \( \omega_r > 0 \). Another option is to impose desired travel times \( \tau_{\text{desired}}(d) \):

\[
j_{\text{DTT}} = \sum_{d=1}^{N} \sum_{r \in R} \omega_r \left( \tau_{\text{route}}(d) - \tau_{\text{desired}}(d) \right)^2
\]

with \( R \) the set of all routes. To prevent large variations in the control input \( c \), a penalty can be formulated for these variations:

\[
j_{\text{var}} = \sum_{d=1}^{N} \left( c(d) - c(d-1) \right)^2.
\]

The total cost function is often a weighted combination of the different costs: 

\[
J = w_1 j_{\text{TT}} + w_2 j_{\text{DTT}} + \cdots + w_{\text{var}} j_{\text{var}}
\]

with \( w_j > 0 \).
into several independent partial queues, one for each route that uses the link. Let us now compute the time spent in each of these partial queues.

For each vertex \( v \) in the network we introduce event times \( t_{v,i}(d) \). Between two events, the variables in the network are constant. An event time can correspond to two types of changes around the vertex:

- a change in the input flow of one of the upstream links connected to the vertex,
- a partial queue becoming empty on one of the upstream links of vertex.

Since these changes are not known beforehand, the traffic is simulated from the current event time \( t_{v,i}(d) \) until the next known event time \( t_{v,i+1}(d) \). When during the simulation of this period one of the two changes appears, a new event time \( t_{v,new}(d) \) is created. The computations then have to be (re-)done for the period \( [t_{v,i}(d), t_{v,new}(d)] \), which leads to re-definition of the time instants with \( t_{v,i+2}(d) := t_{v,i+1}(d) \) and \( t_{v,i+1}(d) := t_{v,new}(d) \).

3.2 Link variables

Consider link \( l \), with its upstream vertex \( v^u_l \) and its downstream vertex \( v^d_l \), as in Fig. 4.

The inflow \( Q^\text{in}_l(d,i) \) of each link \( l \) is given in the timing of the upstream vertex, and present during the period \( [t_{v^u_l,i}(d), t_{v^u_l,i+1}(d)] \), which is the \( i \)th period for the vertex \( v^u_l \). This flow experiences a time delay equal to the free flow travel time, and then becomes the flow that enters the queue in the link \( Q^\text{in}_l\text{queue}(d,i) \) during \( [t_{v^u_l,i}(d), t_{v^u_l,i+1}(d)] \), which is period \( i \) in the timing of the downstream vertex:

\[
Q^\text{in}_l\text{queue}(d,i) = Q^\text{in}_l(d,i)
\]

with \( t_{v^u_l,i}(d) = t_{v^u_l,i}(d) + t^\text{in}\text{delay}(d) \) for all \( i \).

Recall that we assume that in each link there is a partial queue for each route. The average number of vehicles at the start of time period \( i \) in the queue at the end of link \( l \) belonging to route \( r \) is denoted by \( N^\text{veh}_l(d,i) \). As indicated above, the inflow of the queue is given by \( Q^\text{in}_l\text{queue}(d,i) \). The amount of traffic that can leave the queue depends on different factors: (1) the outflow limit \( Q^\text{out}_l(d) \) of the link, (2) the number of queues on the link and their length, (3) the capacity of the downstream links, and (4) the size of the flows that want to enter the downstream links.

We first introduce factors \( \gamma_r(d,i) \) that divide the outflow limit \( Q^\text{out}_l(d) \) proportionally over the different queues (see Fig. 5):

\[
\gamma_r(d,i) = \frac{N^\text{veh}_l(d,i)}{\tau} + Q^\text{in}_l\text{queue}(d,i)
\]

\[
\sum_{r \in R_l} \left( \frac{N^\text{veh}_l(d,i)}{\tau} + Q^\text{in}_l\text{queue}(d,i) \right)
\]

where \( \tau \) is a delay factor representing the time that vehicles require to leave the queue.

![Fig. 4. A link \( l \) with upstream vertex \( v^u_l \) and downstream vertex \( v^d_l \).](image)

![Fig. 5. Routes on a link should share the available outflow](image)

![Fig. 6. Inflow capacity shared by entering flows](image)
Now we can compute the average time spent in the queue, the objective function.

\[ t \text{ time instant} \]

In order to compute the average time the vehicles spend in the queue, we first compute the area under the curve. This is done at the end of the simulation, so that the event timing is completely fixed. We assume that the event timing is large enough and \( q_0 \) is completely fixed. We assume that the time instant \( t_{q_{1}}(d) \) is done at the end of the simulation, so that the event timing of the simulation period \( T \) is re-done.

At this moment the partial queue for the traffic on link \( l \) going via route \( r \) becomes empty, (see Fig. 7). This means that a new time instant \( t_{q_{1}}(d) + T_{r}(d, i) \) should be added to the timing of the downstream vertex \( v_{d} \), and the computations for the current period should be re-done.

After the computations for the whole period are performed, the total number of vehicles in a link can be plotted as in Fig. 8.

3.3 Average time in the queue

In order to compute the average time the vehicles spend in the queue, we first compute the area under the \( N_{q_{r}}^{\text{veh}}(d, \cdot) \) curve. This is done at the end of the simulation, so that the event timing is completely fixed. We assume that \( T \) is large enough and that the demands become small enough towards the end of the period \([0, T]\) so that the queues are completely empty at the end of the simulation\(^1\). If we denote the area under the \( N_{q_{r}}^{\text{veh}}(d, \cdot) \) curve between \( t_{q_{1}}(d) \) and \( t_{q_{1}}(d) + T_{r}(d, i) \) by \( A_{t}(d, i) \), there are two possible cases:

- If \( N_{q_{r}}^{\text{veh}}(d, i) > 0 \) or \( N_{q_{r}}^{\text{veh}}(d, i + 1) > 0 \) then we have
  \[
  A_{t}(d, i) = \frac{1}{2} \left( N_{q_{r}}^{\text{veh}}(d, i) + N_{q_{r}}^{\text{veh}}(d, i + 1) \right) (t_{q_{1}}(d) + T_{r}(d, i) - t_{q_{1}}(d))
  \]

- If \( N_{q_{r}}^{\text{veh}}(d, i) = 0 \) and \( N_{q_{r}}^{\text{veh}}(d, i + 1) = 0 \) we have \( A_{t}(d, i) = 0 \) since the \( N_{q_{r}}^{\text{veh}}(d, \cdot) \) curve is horizontal.

Now we can compute the average time spent in the queue according to

\[ \text{slope: } Q_{q_{r}}^{\text{in queue}}(d, i) - Q_{q_{r}}^{\text{eff}}(d, i) \]

\[ N_{q_{r}}^{\text{veh}}(d, i + 1) = \max \left( 0, N_{q_{r}}^{\text{veh}}(d, i) + (Q_{q_{r}}^{\text{in queue}}(d, i) - Q_{q_{r}}^{\text{eff}}(d, i))(t_{q_{1}}(d) + T_{r}(d, i)) \right). \]

If \( N_{q_{r}}^{\text{veh}}(d, i) + (Q_{q_{r}}^{\text{in queue}}(d, i) - Q_{q_{r}}^{\text{eff}}(d, i))(t_{q_{1}}(d) + T_{r}(d, i)) < 0 \), the queue length already becomes 0 at some time \( t_{q_{1}}(d) + T_{r}(d, i) \) with

\[ T_{r}(d, i) = N_{q_{r}}^{\text{veh}}(d, i)/(Q_{q_{r}}^{\text{in queue}}(d, i) - Q_{q_{r}}^{\text{eff}}(d, i)) \]

\[ \text{where } n_{q_{r}} \text{ is the number of periods in the timing of vertex } v_{r}. \]

3.4 Origin modeling

We model the origin as a virtual link with length 0, see Fig. 9. The demand is given by \( D(d, i) \), with event times \( t_{i} \). This demand is divided over the routes via the turning rates

\[ Q_{o_{r}}^{\text{eff}}(d, i) = \beta_{r}(d) D(d, i) \]

where \( Q_{o_{r}}^{\text{eff}}(d, i) \) is the flow that enters the virtual link during the period \([t_{i}(d), t_{i+1}(d)]\). This flow enters the partial queues that can be present on this link. Further, the origin is modeled in the same way as the links inside the network (see Sections 3.2 and 3.3).

3.5 Turning rates

The turning rates are computed based on the difference in travel time experienced by the drivers. The drivers will change their route when the travel time on another route is shorter:

\[ \zeta_{r}(d + 1) = \max \left( 0, \beta_{r}(d) + \sum_{p \in R} \kappa_{p} \left( \frac{t_{p_{\text{route}}}(d)}{t_{p_{r}}(d) - t_{p_{\text{route}}}(d)} \right) \right). \]

Here \( \kappa_{p} \) includes the fraction of drivers on route \( p \) that change their route from one day to the next based on the travel time difference. Since the sum of the turning rates should be 1, they should be normalized:

\[ \beta_{r}(d + 1) = \sum_{p \in R} \kappa_{p}(d + 1) \]

4. CONTROL APPROACH

In Section 2 we have formulated the control problem, which is actually an optimal control problem. Except when the prediction and control horizons are small, this problem cannot be
solved on-line due to the required computation time. To obtain a good and more robust solution, we determine sub-optimal settings for the control measures with a route choice control approach based on Model Predictive Control (MPC).

4.1 Model predictive control

In MPC for route choice control the goal is to determine the control inputs c that — given the current state of the network, the future demand, and a model of the system — optimize a cost function $J(d)$ over a given prediction period of $N_p$ days ahead, subject to operational and other constraints. This results in a sequence of optimal control inputs $c^*(d), c^*(d+1), \ldots, c^*(d+N_p-1)$. To reduce the computational complexity a control horizon $N_c$ ($N_c < N_p$) is introduced and the control sequence is constrained to vary only for the first $N_c$ days, after which the control inputs are set to stay constant, i.e. $c(j) = c(d+N_c-1)$ for $j = d+N_c, \ldots, d+N_p-1$.

MPC uses a receding horizon approach, i.e. of the optimal control signal sequence only the first sample $c^*(d)$ is applied to the system. Next, at day $d+1$, the procedure is repeated given the new state of the system, and a new optimization is performed for days $d+1$ up to $d+N_p$. Of the resulting control signal again only the first sample is applied, and so on.

In this paper we use real values for the control signal, and as prediction model we use the model described in Section 3. Note however that the proposed approach is generic and modular. So if required other, more complex models could be used instead.

The cost function that is selected for the optimal control problem in Section 2 can now be adapted for the MPC controller. Each day, the model is used to predict the traffic and to compute the cost for the prediction period covering days $d$ up to $d+N_p-1$. For the total travel time this results in:

$$J^{\text{TT}}(d) = \sum_{j=d}^{d+N_p-1} \sum_{r \in R} \beta_r \left( \sum_{i=0}^{n} D(i, j)(t_{i+1} - t_i) \right) \tau_{r \text{route}}(j)$$

The other MPC cost functions can be defined in a similar way.

4.2 Optimization algorithms

MPC uses an optimization algorithm to determine the optimal settings for the control variables. Since the prediction model is a nonlinear model, and since we assume real-valued control inputs, the optimization problem is a nonlinear nonconvex real-valued problem. To solve this type of problems multi-start local search methods (like SQP) and (semi-)global optimization methods (like genetic algorithms, pattern search, or simulated annealing) can be used. Those approaches in principle only yield a suboptimal solution since — in particular for larger networks or longer control horizons — it is in practice not tractable to find the global optimum of the optimization problems that arise in MPC for route choice control. In van den Berg et al. (2008a) we have addressed this problem by considering a limited set of discrete values for the control inputs, by considering a constant demand, and by considering linear cost functions, which allowed us to transform the MPC optimization problem into a mixed integer linear programming (MILP) problem. In van den Berg et al. (2009) the introduction of piecewise constant demands resulted in an optimization problem that can be transformed into an MILP problem under certain assumptions. This approach is still valid for the extension to overlapping routes that is described in this paper.

Now we illustrate the proposed route choice control approach with a simple simulation example involving variable speed limit control. Consider the network with multiple routes given in Fig. 10 with the routes defined as in Fig. 2. This network consists of four links. Link 1 is a long freeway link, with $\ell_1 = 100$ km, has a high capacity, $Q^{\text{cap}} = 6000$ veh/h, and a high maximum speed, $v^{\text{max}}_1 = 120$ km/h. Link 2 is an urban link, with $\ell_2 = 40$ km, $Q^{\text{cap}}_2 = 1000$ veh/h, and $v^{\text{max}}_2 = 50$ km/h. The third link is a freeway link with $\ell_3 = 80$ km, a regular capacity $Q^{\text{cap}}_3 = 4000$ veh/h, and $v^{\text{max}}_3 = 120$ km/h. Link 4 is an arterial road with $\ell_4 = 60$ km, $Q^{\text{cap}}_4 = 2000$ veh/h, and $v^{\text{max}}_4 = 100$ km/h. A driver has to choose between link 1 and 2, and between link 3 and 4. Link 2 has a lower free flow travel time than link 1, and will thus get the preference of the drivers. But due to the low capacity a large queue could be formed, which would increase the travel time and make link 1 more favorable than link 2. Similarly, link 4 has a shorter free-flow travel time than link 3, but due to the lower capacity a longer queue will be formed. In this case it can be expected that the drivers will not a priori prefer one link above the other.

We simulate a period of 15 days, with $T = 120$ min $= 2$ h. The demand pattern is given in Table 1. We have a learning rate of $\kappa = 0.25$ for all routes $r \in R$, a time constant $\tau = 0.33$ h, and the following initial turning rates: $\beta_1(0) = 0.5, \beta_2(0) = 0.1, \beta_3(0) = 0.3$, and $\beta_4(0) = 0.1$. Variable speed limit control is applied, with $v_{\text{min}}^1 = 60$ km/h, $v_{\text{min}}^2 = 15$ km/h, $v_{\text{min}}^3 = 60$ km/h, and $v_{\text{min}}^4 = 30$ km/h. The overall objective of the controller is given by $J = J^{\text{DTT}} + w_J^\text{rel}$, with $J^{\text{DTT}}$ and $J^{\text{rel}}$ defined by (2) and (3) respectively, and with desired travel times $\tau_d(\tau^1 = \tau^2 = \tau^3 = \tau^4 = 2$ h for all days $d$, and weights $w_1 = w_2 = w_3 = w_4 = 1$ and $w = 10^{-5}$. The $J^{\text{DTT}}$ part of the objective function penalizes traffic traveling via routes 3 and 4 (which contain the urban link 2) and favors the others routes so as to divert traffic away from the urban area. In addition, we impose that constraint that the total flow on link 4 should stay below 1750 veh/h in order to prevent too much noise and emissions in the nature reserve along this road. We set $N_0$ to 6 days and $N_0$ to 5 days. The MPC optimization problem is solved using multi-start SQP.

The results are shown in Fig. 11 for the uncontrolled situation and the controlled situation with the constraint on the maximal flow on link 4. The $J^{\text{DTT}}$ value without control is $10.531 \text{h}^2$. If we apply MPC-based control without the constraint for link 4,
we obtain a value of 6.268 h² for $J_{\text{DTT}}$. However, in both cases the maximal flow on link 4 goes as high as 2000 veh/h. If we apply MPC-based control with the constraint for link 4, then the $J_{\text{DTT}}$ value is equal to 10.248 h² and then the flow on link 4 stays below the maximal allowed value of 1750 veh/h.

These results illustrate that MPC can be used to control the travel times on the various routes in a traffic network via dynamic speed limits.

6. CONCLUSIONS

We have developed a control approach based on model predictive control to influence the route choice of drivers, using existing control measures like outflow limits or variable speed limits. In previous work (van den Berg et al. (2008a,b, 2009)) we have considered route choice control for basic networks and constant or piecewise constant demands. In this paper we have extended the model to include networks with overlapping routes and with restricted link inflow capacities. We have used the resulting model in a control approach based on model predictive control (MPC). This approach has been illustrated the control approach with an example, where we applied speed limit control for a simple network with four routes.

Future research will include: extending the approach to multiple origins and destinations, including more traffic control measures, and developing faster optimization algorithms.

acknowledgements

Research supported by the NWO-CONNEKT project “Advanced Multi-Agent Control and Information for Integrated Multi-Class Traffic Networks (AMICI)”, the BSIK projects TRANSUMO and NGI, the European COST Action TU0702, the Transport Research Center Delft, and the Delft Research Center Next Generation Infrastructures.

REFERENCES


E. Cascetta, P. Coppola, and V. Adamo. A day-to-day dynamic model of departure time and pre-trip route choice in presence of advanced travellers information systems. In Proc. 7th World Congres on Intelligent Transportation Systems, Turin, Italy, November 2000.


