Distributed Controller Design for Dynamic Speed Limit Control Against Shock Waves on Freeways

Andrey Popov ∗ Robert Babuška** Andreas Hegyi*** Herbert Werner *

* Institute of Control Systems, Hamburg University of Technology, Germany (andrey.popov@gmx.net, h.werner@tu-harburg.de)
** Delft Center of Systems and Control, TU-Delft, The Netherlands, (r.babuska@tudelft.nl)
*** Transport and Planning, Faculty of Civil Engineering and Geosciences, TU-Delft, The Netherlands, (a.hegyi@tudelft.nl)

Abstract: Shock waves are special types of relatively short traffic jams that propagate opposite to the driving direction. These jams increase travel time, air pollution, and negatively impact safety. One way of dealing with shock waves is to impose dynamic speed limits to eliminate them. Control strategies proposed so far are based either on expert knowledge, or are centralized controllers with high computational demand, such as model predictive control. In this paper, we design decentralized feedback controllers with a fixed structure. For the purpose of design, we use a direct optimization technique capable of dealing with the design objectives and the non-linear character of the system. The advantages of using such simple controllers are that they do not require extensive on-line computations, use only local information and are therefore more attractive from the implementation point of view. We show that a simple, static control law achieves performance similar to previous results with centralized control for the considered scenario. The controller successfully resolves the shock wave and reduces the total time spent by 20%, compared to the uncontrolled case.

1. INTRODUCTION

Wasted time and energy, higher accident risk and increased pollution are just some of the problems caused by traffic congestion on freeways. Much research has recently been focused on solving these problems - from dynamic route guidance, through intelligent vehicles, to building better infrastructure. Dynamic speed limits, already implemented on gantries over freeways in Germany, the Netherlands, USA and several other countries (see, e.g., Lenz et al. [2001], van den Hoogen and Smulders [1994]) appear to be an attractive, easy to implement and cost-effective approach for the future.

In this paper, we focus on the problem of dynamic speed limit control design for a typical freeway stretch. We consider the problem of reducing/resolving shock waves. Such shock waves typically emerge from traffic jams or are triggered in dense traffic by a disturbance, and consist of a relatively short area of high density. The vehicles entering that area are forced to slow down, whereas the vehicles leaving it can speed up again, giving the general effect of a wave propagating in the upstream direction (opposite to the traffic). When not controlled, such moving shock waves can persist for several hours and travel many kilometers (Lenz et al. [2001]).

Control strategies proposed so far are based on model predictive control (MPC) as in Hegyi et al. [2005], or on expert knowledge (Lenz et al. [2001], Smulders [1992]). Here we aim to design a simple, i.e., low-order or static, distributed controller, that uses only local information. Such a solution reduces the cost of communication and the on-line computational load, and increases the modularity and scalability of the system.

The distributed controller design is a difficult problem for which standard optimal and robust design techniques cannot be directly used. The reason is that constraint on the controller structure makes the problem non-convex. Existing alternatives for design either impose special controller structures or rely on assumptions for the underlying system model (see, e.g., Scherer [2000]). Two gradient-based optimization technique proposed by Burke et al. [2006] and Apkarian and Noll [2006] allow fixed-structure $H_{\infty}$ controller design and thus also decentralized controller design, but since they require gradient information they are not directly applicable to other design problems. The work of Dullerud and D'Andrea [1999] views large scale systems as spatially interconnected ones and allows distributed controller design. D'Andrea and Dullerud [2003] extend this to an infinite spatial interconnection of identical subsystems, and reduce the design problem to controller design for only one of the subsystems.

The aforementioned techniques rely on linear system models. For systems like the freeway traffic, where a linear model is difficult to identify, optimization based on a nonlinear model, is often the design method of choice. In the cases when the resulting cost function is non-smooth or has discontinuities there are typically multiple optima and thus a global optimization method is needed.
In this work, the shock wave resolution problem is formulated as a distributed control design optimization problem, where one controller corresponds to one freeway section. Furthermore, because the sections in the model have equal length and parameters, we design a single controller and then replicate it for each section. Due to the model non-linearities and non-quadratic design objectives, numerical simulations are used to evaluate the values of the optimized function. To deal with the resulting non-convex design problem, evolutionary algorithms are chosen as optimization technique.

The paper is structured as follows. In Section 2 the controller design technique is presented. In Section 3 the traffic flow model is introduced. The model parameters, traffic scenario and design objectives are presented in Section 4. In Section 5 the controller structures are introduced. Numerical results are given in Section 6, and Section 7 concludes the paper.

2. DESIGN METHODOLOGY

Like the freeway traffic system, many systems of practical interest are large scale systems and often with distributed parameters. Examples are winding machines, power networks, arrays of micro sensors and actuators, etc. For these problems, the traditional optimal and robust controller design will result in centralized controllers of an order greater or at least equal to the order of the whole system. Such controllers are in most cases impractical for implementation due to the high computational and communication load.

An alternative approach is to use distributed control, where the control task is distributed (often physically) between several controllers, each using local information and possibly communicating with the others. In the case when the controlled system consists of identical subsystems, which is the case for the traffic flow model (introduced in Section 3) one can also choose the distributed controller to have the corresponding structure of identical sub-controllers.

2.1 Distributed Control by Identical Sub-Controllers

Consider a system $G$, that is an interconnection of $N$ identical subsystem $g$. The design task is to find a linear, distributed controller $C(\theta)$, which is an interconnection of $N$ identical sub-controllers $c(\theta)$, where each sub-controller controls one of the subsystems, such that a specified performance index $J$ is optimal, and where $\theta$ are the tuning parameters. Fig. 1 shows an example of a serially interconnected systems and controllers.

![Interconnection of N identical subsystems](image)

If $g$ and $G$ are linear time-invariant systems, one could use the spatially interconnected system representation methodology and reduce the problem to the design of a single $c(\theta)$ (e.g., see D’Andrea and Dullerud [2003]). Note that the methods of Burke et al. [2006] and Apkarian and Noll [2006] are not applicable to this design problem, due to the replicated structure of $c(\theta)$ in $C(\theta)$.

For a non-linear system, none of the above methodologies can be applied. Many systems of practical interest have severe non-linearities, as the traffic system considered here, making it difficult to obtain good linear models. An alternative in these cases is to simulate the closed-loop system with different $c(\theta)$ and evaluate the respective performance criteria. A suitable optimization method is needed to find the optimum of the performance criterion.

The advantages of using simulation-based optimization are:

- Non-linear systems and/or controllers can be handled directly.
- The controller structure can be freely chosen.
- Any design objectives can be used – nonlinear (min), max), non-quadratic, etc.

This comes, however, at a price:

- The design time is often longer, due to the need of numerical simulations.
- There are no or limited guarantees for the system performance and stability for scenarios different than the considered ones in terms of other reference signals, noise, disturbances.

2.2 Optimization

Further, to deal with multiple optima caused by non-smooth or discontinuous cost functions (due to the non-linear model, cost functions or from their combination), a global optimization algorithm is needed. Here we use Evolutionary Algorithm (EA) as an optimization technique.

EAs have gained increasing attention in the last years in many control areas (see Fleming and Purshouse [2002] and the references therein). EAs are parallel, stochastic optimization methods, that do not require gradient information. The parallel character of the search stems from the fact that at each iteration the algorithms work with a set (population) of solutions, not with a single solution. The combination of above properties with evolutionary operators, generating the next population, makes them capable of avoiding getting trapped in local optima.

Because of this parallelism, EAs are well suited for solving multi-objective optimization problems. Such problems often arise in practice, when contradicting objectives have to be satisfied simultaneously (de Fonseca [1995]). As will be shown, in the case of freeway traffic systems such objectives are the requirements of achieving maximal throughput in the freeway and using little control action. One method of dealing with such objectives is to combine them to a single one, by weighting each of them appropriately. This is a tedious task, because the weights are rarely known beforehand and several iterations are necessary. Furthermore, as shown in Das and Dennis [1997] this works only if the set of optimal solutions (Pareto-set) is convex.
Another approach is to use multi-objective EA (MOEA), which provide not a single solution, but a set of Pareto optimal solutions (Pareto-set). This set provides information about the trade-off between the design objectives and allows the designer to make an informed decision on which solution to choose.

3. TRAFFIC SYSTEM MODEL

The simulation model used for the controller design is an extended version of the macroscopic traffic flow model METANET (see Kotsialos et al. [2002]). In this paper a simplified model is used, with no on-ramps and off-ramps, or road junctions. This is a valid assumption for a freeway stretch between two road junctions or for a freeway with small in and out traffic along its length (e.g., small villages). If this assumption does not hold, additional measures, such as ramp metering, may be necessary. The presented approach is general and can also be used to design a controller for such a system. For the sake of brevity, we describe only those parts of the model that are relevant for the controller design.

The model represents a freeway stretch of $N$ segments, each of length $L$ [km] and $\lambda$ number of lanes (see Fig. 2). The vehicles enter the stretch at segment 1 (upstream) and leave it from segment $N$ downstream. The mean vehicle density $\rho_i$ and mean speed $v_i$ for each segment are the state variables and are measured ($i = 1 \ldots N$). Furthermore the dynamic speed limits $U_i$ [km/h], displayed on gantries at the entrance of the segments can be controlled.

The system dynamics are described by two equations. The first one expresses the conservation of vehicles

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L \lambda} (q_{i-1}(k) - q_i(k))$$

where $q_i(k) = \rho_i(k)v_i(k)\lambda$ is the outflow from segment $i$.

The second equation expresses the mean driving speed as a sum of the driving speed at the previous time sample, a relaxation term, a convection term and an anticipation term.

$$v_i(k+1) = v_i(k) + \frac{T}{\tau} (V(\rho_i(k)) - v_i(k))$$

$$+ \frac{T}{L} v_i(k) (v_{i-1}(k) - v_i(k))$$

$$- \eta_i(k) \frac{T}{\tau L} \frac{\rho_{i+1}(k) - \rho_i(k)}{\rho_i(k) + \kappa}$$

where $v_i(k)$ is an intermediate speed variable, $\tau$, $\eta_i(k)$, $\kappa$ are model parameters, adjusted to fit the model to the measurement data, and $\eta_i(k)$ is defined as

$$\eta_i(k) = \begin{cases} \eta_{\text{high}} & \text{if } \rho_{i+1}(k) \geq \rho_i(k) \\ \eta_{\text{flow}} & \text{if } \rho_{i+1}(k) < \rho_i(k) \end{cases}$$

expressing the drivers different anticipation behavior at the head and the tail of a traffic jam (Hegyi et al. [2005]).

The speed is limited by the minimum speed $v_{\text{min}}$.

$$v_i(k+1) = \max(v_i(k+1), v_{\text{min}})$$

The desired speed $V(\rho_i(k))$ describes the speed at which the average driver wants to travel given the traffic density and the displayed speed limit in segment $i$. Here we use the formulation from Hegyi et al. [2005]:

$$V(\rho_i(k)) = \min \left( (1 + \alpha)U_i(k), v_{\text{free}} e^{-\frac{1}{2} \left( \frac{\rho_i(k)}{\rho_{\text{crit}}} \right)^2} \right)$$ (3)

where $v_{\text{free}}$ is the drivers’ desired driving speed on a (nearly) empty freeway without speed limit, $\alpha$ is a compliance term (having a small positive value) expressing the drivers tendency to slightly over speed, $p_{\text{crit}}$ is the critical speed above which traffic becomes congested, and $\alpha$ is a model parameter.

By computing (3) for different road densities one derives the fundamental diagram, as shown in Fig 3. The peak of the curve shows the vehicle density for which flow is the highest. A speed limit effects the diagram in free flow.

![Fig. 3. Fundamental diagram – the relation between traffic density and flow, under a particular speed limit.](image)

The upstream boundary conditions are defined by the user definable traffic demand $d_0(k)$ appearing at origin and by speed profile $v_0(k)$. Origins are modeled with a simple queue model, where $w_0(k)$ is the length of the queue. Because of space limitation details are omitted and the interested reader is referred to Hegyi et al. [2005].

4. DESIGN AND TRAFFIC SCENARIO

In this section the model parameters, the shock wave scenario of interest and the design objectives are presented.

4.1 Model Parameters

We consider a freeway stretch with $N = 20$ segments, where each segment is of $L = 0.5$ km length and with $\lambda = 2$ lanes. The speed limits $U_i$ of the central 10 segments can be controlled. Five segments are left uncontrolled (though the speed and density are measured) at each end to avoid the need of a realistic boundary behavior, and to allow easier observation of the effects of the shock wave and the control action. The measurement and control sampling time is $T = 10$ s.
As a free-flow operating point is chosen $v_i = 108.6 \text{ km/h}$ and $\rho_i = 18 \text{ veh/km/lane}$ for $i = 1 \ldots N$. To make the model realistic, the highest speed limit allowed is 120 km/h (set to be the speed limit for the uncontrolled segments) and the lowest limit that can be imposed is 50 km/h. It should be noted, that the driving speed can also vary outside this range.

Measurement noise is added to both speed and density. The noise has zero mean and standard deviation of $n_v = \pm 1.3 \text{ km/h}$ and $n_\rho = \pm 0.5 \text{ veh/km/lane}$ respectively (approx. of 1% of the operating range). To be able to reproduce and compare the results, the noise signals were generated prior to the controller design.

The remaining model parameters are as follows: $\alpha = 0.1$, $\tau = 10.875 \text{ s}$, $\kappa = 160.594 \text{ veh/km/lane}$, $\rho_{\text{crit}} = 27.8045 \text{ veh/km/lane}$, $a = 2.2615$, $v_{\text{free}} = 128.06 \text{ km/h}$, $v_{\text{min}} = 21.28 \text{ km/h}$, $\eta_{\text{high}} = 138.68 \text{ km}^2/\text{h}$, and $\eta_{\text{flow}} = 66.46 \text{ km}^2/\text{h}$. This model is not based on an existing highway, but nevertheless uses realistic model coefficients, to highlight the advantages of the proposed design technique.

### 4.2 Shock Wave Scenario

In the considered scenario a shock wave enters the freeway stretch from the downstream boundary. The front of the shock wave is linearly increasing (decreasing) from $\rho$ = 18 to 75 veh/km/lane. The shock wave appears between time 0.1 and 0.3 h, as shown in Fig. 4. The upstream boundaries conditions are: $v_0 = 108.6 \text{ [km/h]}$, $\rho_0 = 18 \text{ [veh/km/lane]}$ and $d_0 = 3900 \text{ [veh/h]}$.

![Fig. 4. Shape of the shock wave introduced at the downstream boundary.](image)

### 4.3 Uncontrolled Behavior

An uncontrolled system simulation of 1.5 hours of the average speed, for the presented shock wave scenario, is shown in Fig. 5. The low-speed profile of the shock wave can be clearly seen, propagating through the segments in the direction opposite to the traffic flow. The upstream and downstream boundary conditions are kept constant (except when the shock wave is introduced) to the operating point, corresponding to the free-flow condition. The observed reflection of the tail of the shock wave at the upstream boundary is caused by the constant upstream boundary. A real freeway will not end at the upstream boundary and therefore there will be no such reflection.

### 4.4 Design Objectives

One measure for the system performance is the total time spent (TTS) by all vehicles in the freeway section during the considered time interval. TTS is computed as

$$J_{\text{TTS}} = T \sum_{k=1}^{k_{\text{sim}}} \sum_{i=1}^{N} (\rho_i(k) - \rho_i(0))$$

The TTS for the simulation in Fig. 5 is 735 veh.hours. For comparison the TTS in the case of a free-flow is 549 veh.hours.

Since sharp changes in the imposed speed limits are undesired, from the point of view of safety and increased driver frustration, an additional term has to be added, to penalize changes in the control action:

$$J_{\text{CTRL}} = \sum_{k=1}^{k_{\text{sim}}} \sum_{i=1}^{N} (U_i(k) - U_i(k-1))^2$$

Finally, as continuous variation of speed limit is not applicable and the actually imposed speed limit can take only predetermined values, we round the controller output to the nearest 10 km/h.

### 5. CONTROL STRUCTURE

We consider a decentralized control structure, where the dynamic speed limit of each segment is controlled by a separate controller. Furthermore, because the freeway section is a chain interconnection of identical segments, we impose the same structure on the controller. To achieve this, the search is reduced to a search over the parameters of only one controller, which is then replicated for each controlled segment. This is advantageous because of the reduced number of optimization parameters, and reduced implementation costs.

The controller structure consists of 10 controllers (one for each controlled segment). We perform controller designs, allowing the controllers to use information from a different number of adjacent segments - starting from using only information from the same segment (resulting in a diagonal control structure) to controllers using information from up to $N_d = 5$ downstream segments and up to $N_u = 1$ upstream segment. The reason for allowing further downstream connections is that they can allow the controllers to obtain earlier information about an upcoming shock wave.
and react appropriately. Furthermore, the controllers do not exchange state information. Relaxing this requirement would require special consideration of the controllers at the boundary segments.

The above introduced controller structure is visualized for one of the controllers in Fig. 6, where \( g \) denotes the segments’ discrete time dynamics, \( c \) is the discrete time controller.

\[
\begin{align*}
&\begin{array}{c}
i-1 \\
i \\
i+1 \\
i+2 \\
i+3
\end{array} \\
&\begin{array}{c}
g \\
g \\
g \\
g \\
g
\end{array} \\
&\begin{array}{c}
U_i \\
\rho_i \\
\rho_i \\
\rho_i \\
\rho_i
\end{array}
\end{align*}
\]

Fig. 6. Controller interconnection structure for one sub-controller using information from \( N_u = 1 \) upstream and \( N_d = 3 \) downstream segments.

As controller inputs are taken the differences of the measured speed and density from the operating point, respectively \( \delta v_i(k) = v_i(k) - v_0 \) and \( \delta \rho_i(k) = \rho_i(k) - \rho_0 \). Similarly, the computed controller output is added to an operating point for the speed limit: \( U_i(k) = U_0 + \delta U_i(k) \). Note, that this operating point can be different from the free-flow operating condition.

The presented control structure is used with static and first-order, discrete dynamic controllers. For the static case, the controller equation is

\[
\begin{align*}
\delta U_i(k) &= [d_{u,1} \ldots d_{u,N_u}] V(k) + [d_{\rho,1} \ldots d_{\rho,N_u}] P(k) \\
V(k) &= [\delta v_i-N_u(k) \ldots \delta v_i-N_u(k)]^T \\
P(k) &= [\delta \rho_i-N_u(k) \ldots \delta \rho_i-N_u(k)]^T
\end{align*}
\]

and for the first-order controller the equations are

\[
x(k+1) = ax(k) + [b_{u,1} \ldots b_{u,N_u}] V(k) + [b_{\rho,1} \ldots b_{\rho,N_u}] P(k) \\
\delta U_i(k) = cx(k) + [d_{u,1} \ldots d_{u,N_u}] V(k) + [d_{\rho,1} \ldots d_{\rho,N_u}] P(k)
\]

where \( N_c = 1 + N_u + N_d \) is the number of connected segments, and \( a, b_{u,j}, b_{\rho,j}, c, d_{u,j} \) and \( d_{\rho,j} \) are the controller coefficients (elements of \( \theta \)) to be optimized, \( j = 1 \ldots N_c \).

6. NUMERICAL RESULTS

As already discussed in Section 4.4, the design has to take into account two objectives - TTS and change in control action. To avoid the need of trial-and-error approach in choosing weights we use the MOEA optimization technique, introduced in Section 2.2 (here the algorithm of Zitzler et al. [2001] – SPEA2 is used). For all designs population size of 40 and 200 generations were used.

As an operating point is chosen \( v_0 = 85 \text{ km/h}, \rho_0 = 27 \text{ veh/km/ lane} \) and \( U_0 = 85 \text{ km/h} \), since this is a point at the fundamental diagram (Fig. 3), with equal distance to the lowest and highest speed limits allowed, thus allowing the controller to have the same operating range in both directions.

The points on the Pareto-surfaces, obtained for several control structures, are visualized in Fig. 7. Because of space constraints only some of the more interesting results are shown. To improve visibility the original Pareto-surfaces were trimmed to 20 points. In the legend, \( \text{P} \) stands for static (proportional) controller, \( \text{1st} \) for a first-order one, and the numbers following are in the form \( +N_d/-N_u \). Each controller structure is given a sequential number (in the square brackets), to allow easier references.

Several observations can be made from Fig. 7. First, one can easily see the lower right end of each surface is corresponding to the uncontrolled case - the control cost equals 0 and TTS of 735 veh.hours. Next by comparing the results of the diagonal control strategy [1] to the others, it becomes clear that by using downstream information, significant performance improvements are possible. This is true, even if only information from the immediate neighbor segments is used - controllers [2]. The slightly poorer results achieved with first-order controllers, in comparison to the static ones, are due to the increased number of decision variables (from 12 for controller structure [3] and 26 for [5]), which causes the search to converge to a local optimum. The population size was increased to 100 and the number of generations to 400 to compensate for it.

Although using information from 5 segments ahead provides better results than using only neighbor information, the latter one is more attractive from implementation point of view, while still providing relatively good results. As an example, simulation results for the driving speed and for the speed limit with one of the static controllers [2] (marked by a large square), are shown in Fig. 8 and 9, respectively. The TTS with that controller is 615 veh.hours. As can be clearly seen from the figure, the controller correctly imposes a speed limit upstream from the shock wave, thus reducing the inflow to it and finally resolving it. This is achieved with a reasonably small change in the imposed speed limits. Validating the above results, by simulating with other, randomly generated, noise signals, shows that the controller successfully resolves the shock wave each time, and achieves TTS between 585 and 620 veh.hours. Shock waves with 20% increased duration or peak density were also successfully resolved. A controller with structure [3] was found to reduce the TTS to the range 575 – 610 veh.hours. This provides approximately 20% reduction of the TTS compared to the uncontrolled.
case, which is similar to results with MPC reported in the literature (e.g., Hegyi et al. [2005]).

Fig. 8. Simulation for the driving speed, when using static controller.

Fig. 9. Simulation of the imposed speed limits, when using static controller.

7. CONCLUSIONS AND FUTURE WORKS

In this work decentralized controllers for dynamic speed limits on freeways were designed using optimization-based controller design. The designed controllers have structure corresponding to the serially-interconnected system structure. It was shown that for the considered shock wave scenario a simple linear, static controller, using only immediate neighbor information, successfully resolves the shock wave and reduces the TTS by 20% compared to the uncontrolled case.

Future research plans include extending the analysis to other shock wave scenarios and road demands, simultaneous controller design for a set of shock wave scenarios, as well as controller design for other types of freeway bottlenecks.

REFERENCES


