Hierarchical model-based predictive control for intelligent vehicle highway systems: Regional controllers

L.D. Baskar, B. De Schutter, and J. Hellendoorn

If you want to cite this report, please use the following reference instead:
Hierarchical Model-Based Predictive Control for Intelligent Vehicle Highway Systems: Regional Controllers

L.D. Baskar, B. De Schutter, and J. Hellendoorn

Abstract—We consider a hierarchical control approach for Intelligent Vehicle Highway Systems (IVHS), in which intelligent automated vehicles organized in platoons drive to their destination, controlled by a hierarchical control framework. In this framework there are several levels of controllers starting from on-board controllers inside the vehicles, over platoon controllers, to roadside and area controllers, and finally regional and supraregional controllers. In earlier work we have already described the roadside and the area controllers. In this paper we extend these results by focusing on the higher-level regional and supraregional controllers and by describing in more detail how the interactions between the different control levels work.

I. INTRODUCTION

Traffic congestion is a growing problem throughout the world. There are several possible approaches to reduce the frequency and impact of traffic jams (e.g., building new infrastructure, introducing road pricing policies, stimulating modal shift, promoting public transportation, etc.). On the longer term one of the most promising approaches is the integrated use of traffic management and control systems that incorporate intelligence in both the roadside infrastructure and the vehicles, and that are called Intelligent Vehicle Highway Systems (IVHS).

IVHS are characterized by the presence of fully automated vehicles with throttle, braking, and steering commands being determined by automated on-board controllers. This complete automation of the driving tasks allows organization of the traffic in platoons, i.e., a closely spaced group of vehicles traveling together with short intervehicle distances [1]. Platoons can travel at high speeds and to avoid collisions between platoons at these high speeds, a safe interplatoon distance of about 20–60 m should be maintained. The vehicles within the platoons then travel with small intraplatoon distances of about 2–5 m, which are maintained by the automated on-board speed and distance controllers using Cooperative Adaptive Cruise Control (CACC). By traveling at high speeds while maintaining short intraplatoon distances, the platoon approach allows more vehicles to travel on the network, which improves the traffic throughput [2], [3].

In [4] we have proposed a hierarchical traffic management and control framework for IVHS that builds upon earlier research in this field such as the PATH framework [1]. The control architecture of [4] consists of a multi-level control structure with local controllers at the lowest level and one or more higher supervisory control levels (see also Figure 1). The reason for introducing such a multi-level control hierarchy is that a centralized control solution is not tractable due to the large-scale nature of IVHS. The hierarchical architecture provides a scalable approach where at different levels of the hierarchy different temporal and spatial scales are taken into account.

In previous work we have already proposed model-based predictive control approaches for the roadside controllers [4]–[6] and for the area controllers [7], [8]. In this paper, we will in particular concentrate on the regional and supraregional controllers and discuss a model-based control approach to determine the best division of long-distance flows over the network. We also further integrate our previous results by explicitly focusing on the interfaces between the different levels of the control architecture, in particular on how supraregional, regional, and area controllers should convey commands and instructions to respectively the regional, area, and roadside controllers.

This paper is organized as follows. In Section II we briefly revisit the hierarchical traffic management framework of [4]. Next, the model-based predictive control (MPC) approach we will adopt is briefly recapitulated in Section III. To make the paper self-consistent we also concisely present the main results of [4]–[8] regarding MPC for the roadside and area controllers. In Section IV we then extend MPC to the regional and supraregional controllers. The interfaces between the supraregional, regional, area, and roadside controllers are described in Section V. Section VI concludes the paper.

II. INTELLIGENT VEHICLE HIGHWAY SYSTEMS (IVHS)

The hierarchical control framework for IVHS we have proposed in [4] is based on the platoon concept and distributes the intelligence between the roadside infrastructure and the vehicles using control measures such as intelligent speed adaption, cooperative adaptive cruise control, lane allocation, on-ramp access control, route guidance, etc. The control architecture of [4] consists of a multi-level structure with local controllers on-board the vehicles at the lowest level and one or more higher supervisory control levels, as shown in Figure 1.

The layers of the hierarchical control framework can be characterized as follows:

- The higher-level controllers (such as area, regional, and supraregional controllers) provide network-wide coordination of the lower-level and middle-level controllers.

- In particular, the area controllers provide area-wide...
include a brief recapitulation of the results of [4]–[8].

optimal flows and speeds on links. In this section we give a
for the platoons, and for the area controllers to determine
optimal speeds, lane allocations, and on-ramp release times
(MPC) methods for the roadside controllers to determine
activities (such as maintaining safe intervehicle dis-
platoons, splitting, and lane changing) and intraplatoon
activities (such as maintaining safe intervehicle dis-
ances). The roadside controllers also give instructions for merging, splitting, and lane changes to the platoons.

• The roadside controllers control a part of a highway
or an entire highway. The main tasks of the roadside
controllers are to assign speeds for each platoon, safe
distances (to avoid collisions between platoons), and
release times at the on-ramps. The roadside controllers
do so in their area by providing set-points, reference
trajectories, or control targets. In turn, a group of area
controllers can be supervised by regional controllers, and so on.

• The roadside controllers are responsible for control and
coordination of each vehicle inside the platoon. These
controllers are mainly concerned with actually executing
the interplatoon maneuvers (such as merging with other
platoons, splitting, and lane changing) and intraplatoon
activities (such as maintaining safe intervehicle dis-
tances).

• The vehicle controllers present in each vehicle translate
commands received from the platoon controllers (e.g.,
reference trajectories for speeds (for intelligent speed
adaption), headways (for cooperative adaptive cruise
control), and paths) into control signals for the vehicle
actuators such as throttle, braking, and steering actions.

The lower levels in this hierarchy deal with faster time scales
(typically in the milliseconds range for vehicle controllers up
to the seconds range for roadside controllers), whereas for
the higher-level layers the frequency of updating can range
from few times per minute (for area controllers) to a few
times per hour (for regional and supraregional controllers).

III. MODEL-BASED PREDICTIVE CONTROL FOR THE
ROADSIDE AND AREA CONTROLLERS

In [4]–[8] we have proposed model predictive control
(MPC) methods for the roadside controllers to determine
optimal speeds, lane allocations, and on-ramp release times
for the platoons, and for the area controllers to determine
optimal flows and speeds on links. In this section we give a
short introduction to the general principles of MPC. We also
include a brief recapitulation of the results of [4]–[8].

dynamic route guidance for the platoons, and they su-
ervise and coordinate the activities of the roadside
controllers in their area by providing set-points, reference
trajectories, or control targets. In turn, a group of area
controllers can be supervised by regional controllers, and so on.

• The roadside controllers control a part of a highway
or an entire highway. The main tasks of the roadside
controllers are to assign speeds for each platoon, safe
distances (to avoid collisions between platoons), and
release times at the on-ramps. The roadside controllers
do so in their area by providing set-points, reference
trajectories, or control targets. In turn, a group of area
controllers can be supervised by regional controllers, and so on.

• The roadside controllers are responsible for control and
coordination of each vehicle inside the platoon. These
controllers are mainly concerned with actually executing
the interplatoon maneuvers (such as merging with other
platoons, splitting, and lane changing) and intraplatoon
activities (such as maintaining safe intervehicle dis-
tances). The roadside controllers also give instructions for merging, splitting, and lane changes to the platoons.

• The platoon controllers are responsible for control and
coordination of each vehicle inside the platoon. These
controllers are mainly concerned with actually executing
the interplatoon maneuvers (such as merging with other
platoons, splitting, and lane changing) and intraplatoon
activities (such as maintaining safe intervehicle dis-
tances). The roadside controllers also give instructions for merging, splitting, and lane changes to the platoons.

• The vehicle controllers present in each vehicle translate
commands received from the platoon controllers (e.g.,
reference trajectories for speeds (for intelligent speed
adaption), headways (for cooperative adaptive cruise
control), and paths) into control signals for the vehicle
actuators such as throttle, braking, and steering actions.

The lower levels in this hierarchy deal with faster time scales
(typically in the milliseconds range for vehicle controllers up
to the seconds range for roadside controllers), whereas for
the higher-level layers the frequency of updating can range
from few times per minute (for area controllers) to a few
times per hour (for regional and supraregional controllers).

A. MPC: General principles

As control approach we adopt the model predictive control
(MPC) scheme [9], [10] (see Figure 2). In MPC a discrete-
time model is used to predict the future behavior of the traffic
system (in our case the traffic network or — depending on
the control level — a part of it), and the MPC controller uses
(numerical) optimization to determine the control signals that
result in an optimal system behavior (expressed via, e.g., total
travel time, throughput, emissions, fuel consumption, etc. or
a weighted combination of these) over a given prediction
horizon. The resulting optimal control inputs are applied
using a rolling horizon scheme. More specifically, at each
control step $k$ the state of the traffic system is measured
or estimated, and an optimization is performed over the
prediction period $[kT,(k + N_p)T]$ to determine the optimal
control inputs, where $N_p$ is the prediction horizon. Only the
first value of the resulting control signal (the control signal
for time step $k$) is then applied to the traffic system. At the
next control step $k + 1$ this procedure is repeated.

To reduce the computational complexity and to improve
stability often a control horizon $N_c$ ( $< N_p$) is introduced
in MPC, and after the control horizon has been passed the
control signal is taken to be constant.

There are two loops in MPC: the rolling horizon loop and
the optimization loop inside the controller. The loop inside
the controller of Figure 2 is executed as many times as is
needed to find (sufficiently) optimal control signals at control
step $k$, for the given prediction horizon $N_p$, control horizon
$N_c$, currently measured traffic state, and expected demands.
The loop connecting the controller and the traffic system
is performed once for each control step $k$ and provides the
state feedback to the controller. This feedback is necessary to
correct for the ever present prediction errors, and to provide
disturbance rejection (i.e., compensation for unexpected traf-
cic demand variations). The advantage of this rolling horizon
approach is that it results in an on-line adaptive control
scheme that also allows us to take (slow) changes in the
system or in the system parameters into account by regularly
updating the model of the system.

In MPC the choice of the prediction model is an important
issue. Such a model should provide a balanced trade-off
between accuracy and computational complexity (since the
model will be simulated several times within the optimization
loop; so the simulation speed should be several orders of
magnitude faster than real-time simulation speed). In the next
(sub)sections we will detail which models are appropriate for
the various levels of the control hierarchy.

A second important issue is the right choice of the optimization algorithm. Moreover, sometimes it is possible to approximate the MPC optimization problem by another optimization problem that is much more amenable to fast and efficient optimization such as, e.g., convex or mixed-integer linear optimization problems. The solution of the approximate problem can then be directly applied to the system or it can be used as a good initial solution for the optimization of the original non-linear non-convex optimization problem. These approximation approaches will be also be discussed in the next (sub)sections.

B. MPC for roadside controllers

In order to make the MPC approach tractable the roadside controllers do not consider each individual vehicle in each platoon separately, but they consider each platoon as a single unit and they monitor the movements of the platoons in the highway stretch under their control. More specifically, we model the platoons using the so-called “big car” model, i.e., as a single (long) vehicle with a speed-dependent length. In this way a good trade-off between computational speed and accuracy is obtained.

The control inputs determined for each platoon are its speed, lane assignment, size, as well as release time (at on-ramps) and route choice (at bifurcations). The objective function and constraints can correspond to general traffic performance criteria such as total time spent, throughput, emissions, etc., or they could reflect tracking of targets set forth by the area controllers (see also Section V).

In general, this results in a mixed-integer non-linear optimization problem (if lane allocation and/or size are included in the MPC optimization) or in a real-valued non-linear optimization problem (if lane allocation and size are assigned using heuristics or logic rules). Mixed-integer optimization problems could be solved using genetic algorithms, simulated annealing, or branch-and-bound methods. Continuous optimization problems can be solved using multi-start sequential quadratic programming, genetic algorithms, simulated annealing, or pattern search.

C. MPC for area controllers

In principle, the optimal route choice control problem in IVHS consists in assigning an optimal route to each individual platoon in the network. However, this results in a huge non-linear integer optimization problem with high computational complexity and requirements, making the problem in fact intractable in practice. Since considering each individual platoon is too computationally intensive for on-line real-time control, the area controllers consider a more aggregate model based on flows of platoons. In this context two approaches have been pursued, viz. one based on flow-and-queue model [7] and one based on a METANET-like model for platoons in an IVHS [8].

In the first approach the evolution of the flows (on highway stretches) and queue lengths (at junctions) in the network is described using simple queueing models and assuming a fixed average speed in each highway stretch. The control decisions are then the assignment of flows to the links. Although in general this results in a non-linear, non-convex, and nonsmooth optimization problem, we have shown in [7] that the resulting optimization problem can be approximated using mixed integer linear programming (MILP), for which efficient branch-and-bound solvers are currently available [11].

The second approach is based on a reformulation of the macroscopic traffic flow model METANET [12], [13] for IVHS. The resulting IVHS-METANET model describes the evolution of the traffic flows through average densities, flows, and speeds in the highway segments. The control decisions in this case are the splitting rates at the network nodes and possibly also the speeds on the links. This then results in a non-linear non-convex optimization problem with real-valued variables. To solve the non-linear optimization problem we can use a global or a multi-start local optimization method such as multi-start sequential quadratic programming, pattern search, genetic algorithms, or simulated annealing.

IV. MPC FOR REGIONAL AND SUPRAREGIONAL CONTROLLERS

For the sake of conciseness we will mainly focus on the regional controllers below. However, the proposed approach can also be applied for the supraregional controllers (but then larger) regions have to be considered instead of areas).

In order to keep the MPC optimization problem tractable the regional controllers also have to use an aggregate model to describe the behavior of the traffic system in the areas they control. In this context, we propose the use of an IVHS variant of the Macroscopic Fundamental Diagram.

According to [14], [15] there exists a so-called Macroscopic Fundamental Diagram (MFD) for urban traffic streets describing the relation between space-mean flow and density. The papers [14], [15] also showed that an MFD exists for neighborhood-sized sections of cities, and that the MFD is independent of the demand. Moreover, the outflow of a neighborhood was shown to be more or less proportional to the space-mean flow within the area [14]. This result has been tested in neighborhoods of the order of 10 km² in cities like Yokohama, Japan, based on experimental data gathered by GPS-equipped taxis and fixed detectors. The network MFD obtained is illustrated in Figure 3. It can be used to control network demands to improve accessibility. Simple versions of such control strategies based on the network MFD are already used in London, Stockholm, and Singapore.

Just like we did with the conventional fundamental diagram for human drivers in [8], one could also adopt a modified version of the MFD targeted towards platoons of automated vehicles in an IVHS. The shape of the diagram would then be sharper and the maximal flow would be much higher than in the MFD for human drivers. Together with conservation laws for the densities and the inflows and outflows across areas, the IVHS-MFD can then be used in the prediction model used by the regional controllers. This goes as follows.
A large IVHS network can be divided into several smaller subnetworks corresponding to areas. These subnetworks are controlled and coordinated by the regional controllers so that they typically remain in regions A (free-flow) and B (almost critical) of the MFD (see Figure 3) in order to maximize throughput of the subnetworks. The IVHS network can then be represented by a graph, where the links correspond to the set of areas that feed into node \( \nu \), and the nodes describe the connections between areas, external origins, and/or external exits. Each area or link \( a \) is characterized by an inflow \( q_{\text{in},a}(k) \), an outflow \( q_{\text{out},a}(k) \), and a density \( \rho_a(k) \) at time step \( k \). The external origins \( o \) are characterized by an inflow \( q_{\text{orig},o}(k) \), and the external exits \( e \) are characterized by an outflow \( q_{\text{exit},e}(k) \).

In general the network MFD for IVHS results in a static description of the form\(^1\)

\[
q_{\text{out},a}(k) = \mathcal{M}_a(\rho_a(k)) \quad .
\]

To describe the evolution of the densities inside each area we use a simple conservation equation:

\[
\rho_a(k+1) = \rho_a(k) + \frac{T}{L_a} (q_{\text{in},a}(k) - q_{\text{out},a}(k)) \quad ,
\]

where \( T \) is the sample time step of the system and \( L_a \) is a measure for the total length of the highways and roads in area \( a \).

Finally, for every node \( v \) in the network there is a balance between the inflows and outflows:

\[
\sum_{a \in \mathcal{N}_{v}} q_{\text{out},a}(k) + \sum_{o \in \mathcal{O}_{\text{orig},v}} q_{\text{orig},o}(k) = \sum_{a \in \mathcal{O}_{v}} q_{\text{in},a}(k) + \sum_{e \in \mathcal{O}_{\text{exit},v}} q_{\text{exit},e}(k) \quad ,
\]

where \( \mathcal{N}_{v} \) is the set of areas that feed into node \( v \), \( \mathcal{O}_{v} \) is the set of areas that receive traffic from node \( v \), \( \mathcal{O}_{\text{orig},v} \) is the set of external origins connected to \( v \), and \( \mathcal{O}_{\text{exit},v} \) is the set of external exits connected to \( v \).

Now we can define an optimization problem based on the model (1)–(3) where the aim is to keep the density in all regions below the critical density; this can be done by using a penalty function:

\[
J_{\text{pen}}(k) = \sum_{j=1}^{N_v} \sum_{a \in \mathcal{O}_{v}} \left( \max(0, \rho_a(k+j) - \rho_{\text{crit},a}) \right)^2 \quad ,
\]

where \( \mathcal{O} \) is the set of areas controlled by the given regional controller and \( \rho_{\text{crit},a} \) is the critical density for region \( a \). In addition, we can also minimize the total time spent (TTS) by all vehicles in the region, which can be expressed as

\[
J_{\text{TTS}}(k) = \sum_{j=1}^{N_v} \sum_{a \in \mathcal{O}_{v}} L_a \rho_a(k+j) T \quad .
\]

Both objectives can be combined into one objective function:

\[
J(k) = J_{\text{pen}}(k) + \gamma J_{\text{TTS}}(k)
\]

for some \( \gamma > 0 \). Moreover, we could, e.g., impose constraints on the maximal flows from one area to another.

In general, optimizing \( J(k) \) subject to (1)–(3) results in a non-linear, non-convex optimization problem. To solve this optimization problem we can use a global or a multi-start local optimization method such as multi-start sequential quadratic programming, pattern search, genetic algorithms, or simulated annealing. However, in general the given non-linear, non-convex optimization problem is characterized by huge computational requirements.

Remark 1: The model can be made more complex by introducing maximal capacities for each area. We then have to introduce queues at the external origins for vehicles that cannot enter an area once the maximal capacity is exceeded. Moreover, (2) has to be modified using an if-then-else statement that specifies that the density cannot further increase once the maximal capacity is reached. The evolution of the origin queues can then be described by a simple queuing model. This extended model also results in a non-linear, non-convex optimization problem with real-valued variables.

An approximation based on mixed-integer linear programming

Now we show that we can also approximate (1) using a piecewise-affine (PWA) description (see also [15]) where this is done for the regular network MFD that can be recast as a system of mixed-integer linear equations and inequalities. Since (2) and (3) are linear, this will result in a mixed-integer linear model that serves as an approximation of (1)–(3).

We will need the following lemmas (see [16], [17]).

**Lemma 2:** Consider a function \( f : \mathcal{X} \rightarrow \mathbb{R} \) where \( \mathcal{X} \subset \mathbb{R}^n \) is a bounded set, and let \( M \) and \( m \) be respectively an upper and lower bound of \( f \) on \( \mathcal{X} \). Then \( [f(x) \leq 0] \Leftrightarrow [\delta = 1] \) is equivalent\(^2\) to

\[
\begin{cases} 
  f(x) \leq M(1 - \delta) \\
  f(x) \geq \varepsilon + (m - \varepsilon) \delta 
\end{cases}
\]

\(^2\)The equivalence is a numerical one due to the fact that \( f < 0 \) does not fit the mixed-integer linear programming framework (where only non-strict inequalities are considered). Hence, this inequality is approximated by \( f \leq \varepsilon \) with \( \varepsilon \) a small tolerance, typically the machine precision, where we assume that in practice the case \( 0 < f < \varepsilon \) cannot occur due to the finite number of bits used for representing real numbers on a computer.
Lemma 3: Consider a function \( f : \mathcal{X} \to \mathbb{R} \) where \( \mathcal{X} \subset \mathbb{R}^n \) is a bounded set, and let \( M \) and \( m \) be respectively an upper and lower bound of \( f \) on \( \mathcal{X} \). Then \( y = \delta f(x) \) with \( \delta \in \{0,1\} \) is equivalent to
\[
\begin{align*}
y &\leq M \delta \\
y &\geq m \delta \\
y &\leq f(x) - m(1 - \delta) \\
y &\geq f(x) - M(1 - \delta)
\end{align*}
\]

To transform (1) into a PWA equation, the range of allowed densities \([0, \rho_{\text{max},a}]\) for area \( a \) is split into \( N_a \) intervals \([\rho_{a,i}, \rho_{a,i+1}]\) with \( \rho_{a,1} = 0 \) and \( \rho_{a,N_a+1} = \rho_{\text{max},a} \). Then (1) can be approximated as
\[
q_{\text{out},a}(k) = \alpha_{a,i} \rho_{a}(k) + \beta_{a,i} \quad \text{if} \quad \rho_{a}(k) \in [\rho_{a,i}, \rho_{a,i+1}]
\]
where \( \alpha_{a,i} \) and \( \beta_{a,i} \) for \( i = 1, \ldots, N_a \) are determined so as to get a close match between the PWA description (5) and the original description (1) (see also [15]). Next we introduce binary variables \( \delta_{a,i}(k) \) such that
\[
\delta_{a,i}(k) = 1 \quad \text{if and only if} \quad \rho_{a,i} \leq \rho_{a}(k) \leq \rho_{a,i+1}.
\]
Using Lemma 2 this constraint can be transformed into a system of mixed-integer linear inequalities.

Now (5) can be rewritten as
\[
q_{\text{out},a}(k) = \sum_{i=1}^{N_a} \alpha_{a,i} \delta_{a,i}(k)
\]
By introducing real-valued auxiliary variables \( y_{a,i}(k) = \rho_{a}(k) \delta_{a,i}(k) \) (which can be written as a system of mixed-integer linear inequalities by Lemma 3), (7) can be written as a linear equation:
\[
q_{\text{out},a}(k) = \sum_{i=1}^{N_a} \alpha_{a,i} y_{a,i}(k) + \beta_{a,i} \delta_{a,i}(k)
\]
If we combine all equations and inequalities, we obtain a system of mixed-integer linear inequalities.

Remark 4: Since one and only one \( \delta_{a,i}(k) \) is equal to 1 at the same time, we have \( \sum_{i=1}^{N_a} \delta_{a,i}(k) = 1 \). So we can eliminate \( \delta_{a,N_a}(k) = 1 - \sum_{i=1}^{N_a-1} \delta_{a,i}(k) \), thereby reducing the number of binary variables by \(|\mathcal{A}| N_p\) where \(|\mathcal{A}|\) denotes the number of areas, while still obtaining a system of mixed-integer linear inequalities.

Let us now consider the objective function \( J(k) \). Note that \( J_{\text{TTS}}(k) \) defined above is linear, but \( J_{\text{pen}}(k) \) is not. However, if we remove the square in (4), we get a PWA objective function. By introducing some dummy variables and by using an approach that is similar to the one above, and by subsequently including the mixed-integer linear inequalities obtained above, the original MPC optimization problem can be recast into a mixed-integer linear programming (MILP) problem of the form
\[
\begin{align*}
\min \quad & c^T x \\
\text{s.t.} \quad & Ax \leq b \\
& Cx = d
\end{align*}
\]
for appropriately defined matrices and vectors \( A, b, c, C, d \), where \( x \) contains all variables for MPC step \( k \). For the above problem efficient algorithms have been developed [11].

The solution of the MILP problem can then be directly applied to the system, or it can be used as a good initial starting point for an optimization run for the original non-linear, non-convex MPC optimization problem.

A similar approach can also be used for the extended model mentioned in Remark 1.

V. INTERFACES BETWEEN THE DIFFERENT CONTROL LAYERS

The higher-level controllers can influence the controller in the level immediately below them in two ways: by specifying weights, set-points, or reference signals in the objective function, or by specifying targets or thresholds in the constraints. The lower-level controller then has to solve an optimization problem of the form
\[
\begin{align*}
\min_{u(k), \ldots, u(k+N_p-1)} & \quad J(k) = J_{\text{high}}(k) + \lambda J_{\text{local}}(k) \\
\text{s.t.} \quad & x(k+j+1) = f(x(k+j), u(k+j), d(k+j)) \\
& \quad \text{for } j = 0, \ldots, N_p - 1 \\
& u(k+j) = u(k+N_c - 1) \quad \text{for } j = N_c, \ldots, N_p - 1 \\
& C_{\text{high}}(x(k), \ldots, x(k+N_p), u(k), \ldots, u(k+N_c - 1)) \\
& C_{\text{local}}(x(k), \ldots, x(k+N_p), u(k), \ldots, u(k+N_c - 1))
\end{align*}
\]
where \( J_{\text{high}} \) and \( C_{\text{high}} \) represent respectively the objectives and constraints (in the form of a system of equations and/or inequalities) imposed by the higher-level controller, \( J_{\text{local}} \) is the local, additional objectives that have to be optimized, \( \lambda > 0 \) is a weighting factor, \( C_{\text{local}} \) contains the local constraints, and \( x(k+j) \) is the prediction of the state of the traffic system (region, area, highway stretch, depending on the control level) at time step \( k+j \), \( u(k+j) \) is the control input at time step \( k+j \), and \( d(k+j) \) is the estimate of the traffic demand at time step \( k+j \). In addition, the model equations (10) and the control horizon constraint (11) are also included.

In the next subsections we discuss the nature of \( J_{\text{high}} \) and \( C_{\text{high}} \) in more detail for the various levels of the control hierarchy.

A. Supraregional to regional

The supraregional controllers determine flows between regions. The regional controllers could take these into account by including them in \( J_{\text{high}} \) as set-points or reference trajectories to be tracked, and by penalizing deviations from the set-points or reference trajectories using the \( \ell_2 \) or \( \ell_1 \) norm. Alternatively, the flows determined by the supraregional controllers could be included through \( C_{\text{high}} \) by specifying a bandwidth around the prescribed flows, a bandwidth to which the actual flows should be restricted.

B. Regional to area

The interface between the regional controllers and the area controllers can be set up in a similar way as described in Section V-A above.
C. Area to roadside

The control variables determined by the area controllers are the flows on the links and/or the splitting rates at the nodes with more than one outgoing link (and if speed limits are included, also these speed limits). Once the optimal flows, splitting rates, and speeds have been determined by the area controller, they are sent to the lower-level roadside controllers, which can then translate them into actual speed, route, and lane allocation instructions for the platoons. So in this case the communication goes through the performance criterion $J_{\text{high}}$.

In our approach, the roadside controllers can provide lane allocation commands and speeds in order to realize the target flows and speeds in the links. These control measures can then slow down or speed up platoons in the links and also steer the platoons in certain directions depending on the imposed splitting rates for the flows. At the nodes, the roadside controller will additionally provide routing instructions for every platoon on the stretch under its supervision. The roadside controller will determine these routing instructions by taking into account the destinations of the platoons and also the imposed splitting rates or the target flows on the adjacent highways.

The roadside controller can combine the speed and route guidance control measures along with on-ramp access timing to control the platoons that enter from on-ramps. The platoon length will play a crucial role while providing routing instructions to the platoons at internal nodes or bifurcation junctions. So if necessary, the roadside controllers can then also provide commands for platoon splits and merges, and determine new platoon compositions and platoon lengths.

VI. CONCLUSIONS

We have considered a hierarchical control framework for intelligent vehicles highway systems (IVHS). The medium and upper layers of this architecture, viz. roadside, area, and (supra)regional controllers can use model-based predictive control (MPC) to determine the most appropriate control signals for the platoons, flows, areas, or regions under their control. In previous work we have already presented MPC for the roadside and area controllers. The current paper extends this work by proposing an MPC approach for the regional and supraregional controllers and by specifying the interaction between the roadside, area, and (supra)regional controllers.

In our future research, we will consider extensive integrated case studies and assess the performance improvement of the proposed approach with respect to other approaches for the operational management of large-scale IVHS [1], which are mostly either decentralized or heuristic.

ACKNOWLEDGMENTS

We would like to thank prof. Serge Hoogendoorn of the Department of Transport & Planning of Delft University of Technology for drawing our attention to the network Macroscopic Fundamental Diagram.

Research funded by the European 7th framework STREP project “Hierarchical and distributed model predictive control (HD-MPC)”, contract number INFSO-ICT-223854, the European 7th Framework Network of Excellence “Highly-complex and networked control systems (HYCON2)”, the European COST Action TU0702, the BSIK project “Next Generation Infrastructures (NGI)”, the Transport Research Centre Delft, and the Delft Research Center Next Generation Infrastructures.

REFERENCES