A convex formulation for optimal distribution of congestion in multi-region cities

Author(s):
Kouvelas, Anastasios; Saeedmanesh, Mohammadreza; Geroliminis, Nikolas

Publication Date:
2016

Permanent Link:
https://doi.org/10.3929/ethz-b-000276028

Rights / License:
In Copyright - Non-Commercial Use Permitted
A convex formulation for optimal distribution of congestion in multi-region cities

Anastasios Kouvelas *, Mohammadreza Saeedmanesh *, Nikolas Geroliminis *

*Urban Transport Systems Laboratory, School of Architecture, Civil & Environmental Engineering
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
{tasos.kouvelas,mohammadreza.saeedmanesh,nikolas.geroliminis}@epfl.ch

March, 2016

Abstract

Most cities around the world become persistently denser and wider over the last decades and the problem of urban traffic management is steadily gaining momentum due to its social and environmental impacts. Many efforts have been carried out to optimize the signal settings during the peak hours, where networks face serious congestion problems and the performance of the infrastructure degrades significantly. The state-of-practice strategies fail to deal efficiently with oversaturated conditions (i.e. queue spillbacks and partial gridlocks), as they are either designed by use of simplified models that do not accurately replicate the propagation of congestion, or they are based on application-specific heuristics. An alternative approach for real-time network-wide control that has recently gained a lot of interest is the perimeter flow control (or gating). The basic concept of such an approach is to partition heterogeneous cities into a small number of homogeneous regions (zones) and apply perimeter control to the inter-regional flows along the boundaries between regions. The inter-transferring flows are controlled at the intersections located at the borders between regions, so as to distribute the congestion in an optimal way and minimize the total delay of the system. This can be viewed as a high-level control scheme and be combined with other strategies (e.g. local or coordinated controllers) in a hierarchical control framework (this topic has gained a lot of attraction in the research community lately). For a recent review on this research direction the reader is referred to Keyvan-Ekbatani et al. (2012), Ramezani et al. (2015) and Kouvelas et al. (2015). A detailed literature review will be provided in the full paper.

In the current work we focus on the same problem described above and we study a convex formulation of an optimal control problem. The original model for the dynamics of the multi-region process (plant) is highly nonlinear and the modelling tool is the Macroscopic Fundamental Diagram (MFD). MFD provides a concave, low-scatter relationship between network vehicle accumulations [veh] or density [veh/km] and network circulating flow [veh/h] or production [veh·km] for every region of the system. The proposed methodology includes the real-time estimation of some model parameters from measurements and the inference of a simple prediction model from real data. The problem is solved in a rolling optimization horizon, model predictive control (MPC) framework and applied to the nonlinear plant. Different objective functions are investigated and the efficiency of the control decisions is compared to the “ideal” case where the nonlinear MPC problem is solved. Note that this “ideal” approach is more challenging in real life due to lack of data, and most importantly, its computational requirements make the real-time applicability cumbersome.

Consider an urban network partitioned in $N$ homogeneous regions with well-defined MFDs. The index $i \in \mathcal{N} = \{1, 2, \ldots, N\}$ denotes the region of the system, $n_i(t)$ the total accumulation (number of vehicles) in region $i$ and $n_{ij}(t)$ the number of vehicles in region $i$ with final destination region $j \in \mathcal{N}$.
at a given time $t$. Let $N_i$ be the set of all regions that are directly reachable from the borders of region $i$, i.e. adjacent regions to region $i$. The discrete time MFD dynamics of the $N$-region system can be described by the following first order difference equations:

$$n_{ii}(k_m + 1) = n_{ii}(k_m) + T_m \left( q_{ii}(k_m) - M_{ii}(k_m) - \sum_{h \in N_i} M^h_{ij}(k_m) + \sum_{h \in N_i} M^i_{hi}(k_m) \right)$$  \hspace{1cm} (1)

$$n_{ij}(k_m + 1) = n_{ij}(k_m) + T_m \left( q_{ij}(k_m) - \sum_{h \in N_i} M^h_{ij}(k_m) + \sum_{h \in N_i} M^i_{ji}(k_m) \right) \hspace{1cm} i \neq j$$  \hspace{1cm} (2)

where $k_m = 0, 1, \ldots, K_m - 1$ is the discrete time index, $T_m \,[\text{sec}]$ the sample time period of the model (i.e. time $t = k_m T$) and $n_{ii}(k_m) = \sum_{j \in N} n_{ij}(k_m)$ the total accumulation of region $i$. The exogenous variables $q_{ij}(k_m) \,[\text{veh/sec}]$ denote the (uncontrollable) traffic flow demand that is generated in region $i$ at time step $k_m$ with final destination in region $j$ (i.e. $q_{ii}(k_m)$ is the demand generated in region $i$ that has final destination in region $i$). The variables $M^h_{ij}(k_m) \,[\text{veh/sec}]$ denote the transfer flows from region $i$ to region $h$ that have final destination region $j$, while $M_{ii}(k_m) \,[\text{veh/sec}]$ is the internal trip completion rate of region $i$.

We assume that for each region $i$ there exists a production MFD between accumulation $n_i(k_m)$ and total production $P_i(n_i(k_m)) \,[\text{veh·m/sec}]$, which describes the performance of the system in an aggregated way. This MFD can be easily estimated using measurements from loop detectors and/or GPS trajectories. Transfer flows and internal trip completion rates are estimated corresponding to the production MFD of the region and proportionally to the accumulations $n_{ij}(k_m)$ as follows:

$$M_{ii}(k_m) = \theta_{ii}(k_m) \frac{n_{ii}(k_m) P_i(n_i(k_m))}{L_i} \hspace{1cm} (3)$$

$$M^h_{ij}(k_m) = \min \left\{ C^h_{ij}(n_h(k_m)) u_{ih}(k_m) \theta^h_{ij}(k_m) \frac{n_{ij}(k_m) P_i(n_i(k_m))}{L_i} \right\} \hspace{1cm} (4)$$

where $L_i$ is the average trip length for region $i$, which is assumed to be independent of time and destination, internal or external, in $i$. The parameters $\theta_{ii}(k_m), \theta^h_{ij}(k_m)$ are exogenous (e.g. constant) and reflect the route choice. The transfer flows $M^h_{ij}(k_m)$ are the minimum between the sending flow from region $i$ (which only depends on the accumulations of the region), and the receiving capacity $C^h_{ij}(n_h(k_m)) \,[\text{veh/sec}]$ of region $h$. This flow capacity is a function of the accumulation $n_h(k_m)$ and is introduced to prevent overflow phenomena within the regions, i.e. each region $i$ has a maximum accumulation $n_{i,max}$:

$$0 \leq n_i(k_m) \leq n_{i,max}, \forall i \in N$$  \hspace{1cm} (5)

If $n_i(k_m) = n_{i,max}$ the region reaches gridlock and all the inflows along the periphery are restricted. Finally, the control variables $u_{ih}(k_m), \forall i \in N, h \in N_i$ denote the fraction of the flow that is allowed to transfer from region $i$ to region $h$ at time $k_m$. The values of the control variables are physically constrained as follows:

$$0 \leq u_{ih}(k_m) \leq 1, \hspace{1cm} \forall i \in N, h \in N_i$$  \hspace{1cm} (6)

Equations (1)–(4) are a discretized version of equations presented in Yildirimoglu et al. (2015) and represent the traffic dynamics in an $N$-region urban network considering the heterogeneity effect and integrating an aggregated routing model. Note, that these equations assume that drivers can choose any arbitrary sequence of regions as their path and cross region boundaries multiple times.

In order to derive a convex model that can be utilized for real-time control purposes we make the following assumptions:
we introduce the model parameters $\alpha_{ij} = n_{ij}/n_i$. These parameters can be estimated in real-time from measurements (e.g. using Kalman filter or maximum likelihood approximation) and the goal is to develop a simple model that can predict their future dynamics; this can be done with machine learning techniques and historical data

• new “dummy” control variables $u_{ii}$ are introduced that restrict the trip completion rates at every region $i$. Although these variables are not reasonable from a physical point of view they are required in order for the problem to be convex. A conjecture is that the solution of MPC will always result in $u_{ii} = 1, \forall i$, but this needs to be validated with the results

• we approximate the MFDs of the regions with piecewise affine (PWA) functions $G_i(n_i)$ that form a convex set (see Figure 1 for a case study with 4 regions)

• we introduce new decision variables $f_{ih} = u_{ih} \frac{P_i(n_i(k_m))}{L_i}$ and $f_{ii} = u_{ii} \frac{P_i(n_i(k_m))}{L_i}$, $i \in \mathcal{N}, h \in \mathcal{N}_i$

• we drop the capacity constraint in equation (4); this is a reasonable simplification as MPC solutions will not let the system to operate close to gridlock

The assumptions outlined above are reasonable approximations/simplification of the nonlinear plant in order to derive a convex formulation that can be used for online MPC. The derived convex optimization problem that approximates the original system and can be solved online is as follows

$$\min_{f_{ij}(k_e), n_i(k_e)} \sum_{k=0}^{K_e-1} J(n(k))$$
subject to  
\[ n_i(k_c + 1) = n_i(k_c) + \]
\[ T_c \left( q_i(k_c) - \sum_{j \in \mathcal{N}_i} \alpha_{ij}(k_c) f_{ij}(k_c) + \sum_{j \in \mathcal{N}_i} \alpha_{ji}(k_c) f_{ji}(k_c) - \left( 1 - \sum_{j \in \mathcal{N}_i} \alpha_{ij} \right) f_{ii}(k_c) \right) \]  (8)
\[ 0 \leq f_{ij}(k_c) \leq G_i(k_c) \]  (9)
\[ 0 \leq f_{ii}(k_c) \leq G_i(k_c) \]  (10)
\[ 0 \leq n_i(k_c) \leq n_i,_{\text{max}} \]  (11)
\[ k_c = 0, 1, \ldots, K_c - 1, \ i \in \mathcal{N}, \ j \in \mathcal{N}_i \]  (12)

where \( n = \text{vec}(n_i) \) and \( J(n) \) any convex function (e.g. \( J = \sum_{k=0}^{K_c-1} \sum_{i \in \mathcal{N}} n_i(k) \) represents the total delay, but also a quadratic function can be used, e.g. \( J = \sum_{k=0}^{K_c-1} n^\top(k) Q n(k) \), where \( Q \) is an appropriate weighting matrix). Note that all the constraints of this problem are linear and as a consequence the computational requirements are quite low, even for a network with many regions and large prediction horizons.

Figure 2 presents a demand scenario for the case study with 4 regions (e.g. 4x4 OD matrix). For this demand Figure 3 displays the evolution of accumulations for the plant and for the model used for MPC. The prediction horizon of the linear model is 10 times higher than the sample time of the plant (e.g. \( K_c = 10 k_m \)). The trajectories of the accumulations demonstrate that this model can be used to approximate the original nonlinear one for small prediction horizons, thus it is appropriate for the MPC framework. Ongoing and future work deals with investigations about different objective functions for the MPC, estimation and prediction of model parameters \( \alpha_{ij} \), sensitivity of the approach to different OD demand patterns and robustness of MPC to uncertainties.

**Keywords**
Macroscopic fundamental diagram; convex optimization; model predictive control; congestion management.

**References**


Figure 2: Traffic demand for the four regions and all simulation horizon (4×4 OD matrix, \(i\) refers to origin and \(j\) to destination).

Figure 3: Vehicle accumulations for the plant (solid lines) and the linear model (dashed lines) when applied for 10 cycles of prediction (\(K_c = 10k_m\)).