Combined estimation and control of large-scale urban road networks:
A real-time optimization based approach

Extended abstract for hEART 2017, Haifa

Isik Ilber Sirmatel and Nikolas Geroliminis
Urban Transport Systems Laboratory (LUTS)
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

INTRODUCTION
Modeling, estimation, and control of large-scale urban traffic networks present considerable challenges. Inadequate infrastructure and coordination, low sensor coverage, spatiotemporal propagation of congestion, and the uncertainty in traveler choices contribute to the difficulties faced when creating realistic models and designing effective traffic estimation and control schemes for urban networks. Although considerable research has been directed towards designing efficient real-time traffic management schemes in the last decades, estimation and control of heterogeneously congested large-scale urban networks remains a challenging problem.

Traffic modeling and control studies for urban networks usually focus on microscopic models keeping track of link-level traffic dynamics with control strategies using local information. Based on the linear-quadratic regulator (LQR) problem, traffic-responsive urban control (TUC) (Diakaki et al., 2002) and its extensions (Aboudolas et al., 2010; Kouvelas et al., 2011) represent a multivariable feedback regulator approach for network-wide urban traffic control. Although TUC can deal with oversaturated conditions via minimizing and balancing the relative occupancies of network links, it may not be optimal for heterogeneous networks with multiple pockets of congestion. Inspired by the max pressure routing scheme for wireless networks, many local traffic control schemes have been proposed for networks of signalized intersections (see Kouvelas et al. (2014); Varaiya (2013); Wongpiromsarn et al. (2012); Zaidi et al. (2015)), which involve evaluations at each intersection requiring information exclusively from adjacent links. Although the high accuracy of microscopic traffic models is desirable for simulation purposes, the increased model complexity results in complications for control, whereas local control strategies might not be able to operate properly under heavily congested conditions, as they do not protect the congested regions upstream. Another disadvantage of sophisticated local controllers is that they might require detailed information on traffic states, which are difficult to estimate or measure.

Literature on traffic state estimation mainly focuses on freeway networks: A mixture Kalman filter based on the cell transmission model (Daganzo, 1995) is proposed in Sun et al. (2003). In Wang and Papageorgiou (2005), an extended Kalman filter is proposed for real-time state and parameter estimation for a freeway network, the dynamics of which is described by the METANET model (Messner and Papageorgiou, 1990). Mihaylova et al. (2007) develops a particle filtering framework for a second order freeway traffic model that is efficiently parallelizable. Yuan et al.
(2012) reports the superiority of Lagrangian state estimation formulations over the Eulerian case using extended Kalman filters for the Lighthill-Whitham and Richards (LWR) model. There is also some literature on urban traffic state estimation: Pueboobpaphan and Nakatsuji (2006) design an unscented Kalman filter based on a kinematic wave model modified for urban traffic. A combined approach via integrating the Kalman filter with advanced data fusion techniques is taken by Kong et al. (2009) for urban network state estimation. Nantes et al. (2016) propose a data fusion based extended Kalman filter for urban corridors based on the LWR model. Interestingly, even though there is considerable literature on model-based state estimation for freeways, there are very few works on comparable techniques for urban road networks.

An alternative to local real-time traffic control methods is the two layer hierarchical control approach. At the upper layer, the network-level controller optimizes network performance via regulating macroscopic traffic flows through interregional actuation systems (e.g., perimeter control), whereas at the lower layer the local controllers regulate microscopic traffic movements through intraregional actuation systems (e.g., signalized intersections). The macroscopic fundamental diagram (MFD) of urban traffic is a modeling tool for developing low complexity aggregated dynamic models of urban networks, which are required for the design of efficient network-level control schemes for the upper layer. It is possible to model an urban region with roughly homogeneous accumulation (i.e., small spatial link density heterogeneity) with an MFD, which provides a unimodal, low-scatter, and demand-insensitive relationship between accumulation and trip completion flow (Geroliminis and Daganzo, 2008).

The concept of MFD with an optimum accumulation was first proposed by Godfrey (1969), and its existence was recently verified with dynamic features and real data by Geroliminis and Daganzo (2008). Control strategies based on MFD modeling and using perimeter control type actuation (i.e., manipulating transfer flows between neighboring regions) have been proposed by many researchers for single-region (Daganzo, 2007; Gayah et al., 2014; Haddad and Shraiber, 2014; Keyvan-Ekbatani et al., 2012) and multi-region (Aboudolas and Geroliminis, 2013; Haddad and Geroliminis, 2012) urban areas. Application of the MPC technique to the control of urban networks with MFD modeling also attracted recent interest. Geroliminis et al. (2013) design a nonlinear MPC for a simple two-region urban network equipped with a perimeter control system. Haddad et al. (2013) develop an MPC scheme for the cooperative control of a mixed transportation network consisting of a freeway and two urban regions. Hajiahmadi et al. (2015) generalize the two-region MFD network model of Geroliminis et al. (2013) to that of an $R$-region network, and propose hybrid MPC schemes for an urban network equipped with both perimeter control systems and switching signal timing plans. Ramezani et al. (2015) develop a model capturing the dynamics of heterogeneity and design a hierarchical control system with MPC on the upper level. More detailed literature reviews in MFD-based modeling and control can be found in Saberi and Mahmassani (2012) and Yildirimoglu et al. (2015).

Although there is considerable literature on traffic estimation (especially for freeway networks), combined estimation and control for heterogeneously congested large-scale urban networks remains an open problem. In this paper we propose integrated schemes for real-time optimization based estimation and control for urban networks with MFD-based modeling.

MODELING OF URBAN NETWORKS

Consider a heterogeneous urban road network $\mathcal{R}$ that can be partitioned into 2 homogeneous regions (see figure 1), i.e., $\mathcal{R} = \{1, \ldots, n_R\}$ with $n_R = 2$. Each region has a well-defined outflow
MFD, defined via $G_I(N_I(t))$ (veh/s), which is the outflow at accumulation $N_I(t)$. The demand for trips in region $I$ with destination $J$ is $Q_{IJ}(t)$ (veh/s), whereas $N_{IJ}(t)$ (veh) is the accumulation in region $I$ with destination $J$, and $N_I(t)$ (veh) is the total accumulation in region $I$, at time $t$; $I, J \in \mathcal{R}$; $N_I(t) = \sum_{J \in \mathcal{R}} N_{IJ}(t)$. Between the two regions 1 and 2 there exists perimeter controls $U_{12}(t)$ and $U_{21}(t) \in [0, 1]$, that can manipulate the transfer flows. The dynamics of the 2-region MFDs network is (Geroliminis et al., 2013):

$$
\begin{align*}
\dot{N}_{11}(t) &= Q_{11}(t) + U_{21}(t) \cdot M_{21}(t) - M_{11}(t) \\
\dot{N}_{12}(t) &= Q_{12}(t) - U_{12}(t) \cdot M_{12}(t) \\
\dot{N}_{21}(t) &= Q_{21}(t) - U_{21}(t) \cdot M_{21}(t) \\
\dot{N}_{22}(t) &= Q_{22}(t) + U_{12}(t) \cdot M_{12}(t) - M_{22}(t),
\end{align*}
$$

where the $M_{II}(t)$ and $M_{IJ}(t)$ terms express the exit and transfer flows, which can be expressed as follows:

$$
\begin{align*}
M_{II}(t) &= \frac{N_{II}(t)}{N_I(t)} \cdot G_I(N_I(t)) \quad \forall I \in \mathcal{R} \\
M_{IJ}(t) &= \frac{N_{IJ}(t)}{N_I(t)} \cdot G_I(N_I(t)) \quad \forall I \in \mathcal{R}, \quad J \in \mathcal{R} \setminus \{I\}.
\end{align*}
$$

All trips inside a region are assumed to have similar trip lengths (i.e., the origin and destination of the trip does not affect the distance traveled by a vehicle). Simulation and empirical results (Geroliminis and Daganzo, 2008) suggest the possibility of approximating the MFD by an asymmetric unimodal curve skewed to the right (i.e., the critical accumulation $N_I^{cr}$, for which $G_I(N_I(t))$ is at maximum, is less than half of the jam accumulation $N_I^{jam}$ that puts the region in gridlock). Thus, $G_I(N_I(t))$ can be expressed using a third-order polynomial in $N_I(t)$:

$$
G_I(N_I(t)) = A_I N_I^3(t) + B_I N_I^2(t) + C_I N_I(t),
$$

where $A_I$, $B_I$, and $C_I$ are estimated parameters.

The inflow demand terms $Q_{IJ}(t)$ are assumed to have uncertainty, which we model as follows:

$$
Q_{IJ}(t) = \max(D_{IJ}(t) + W_{IJ}(t), 0) \quad \forall I, J \in \mathcal{R},
$$

where $D_{IJ}(t)$ expresses a known average demand profile and $W_{IJ}(t)$ is the associated demand noise with zero mean and normal distribution, i.e., $W_{IJ}(t) \sim N(0, \sigma_W^2)$, with $\sigma_W^2$ specifying the
demand noise variance. Furthermore, it is assumed that the network is equipped with sensors that can measure the accumulations $N_{IJ}(t)$, for which the measurement noise can be similarly modeled as follows:

$$Y_{IJ}(t) = \max(N_{IJ}(t) + V_{IJ}(t), 0) \quad \forall I, J \in \mathcal{R},$$  

where $Y_{IJ}(t)$ is the measurement on $N_{IJ}(t)$ whereas $V_{IJ}(t)$ is the associated measurement noise with zero mean and normal distribution, i.e., $V_{IJ}(t) \sim \mathcal{N}(0, \sigma_V^2)$, with $\sigma_V^2$ specifying the measurement noise variance.

**COMBINED REAL-TIME OPTIMIZATION BASED ESTIMATION AND CONTROL OF LARGE-SCALE URBAN NETWORKS**

**Moving Horizon Estimation**

We formulate the problem of finding the $N_{IJ}$ and $W_{IJ}$ values that minimize the discrepancy between measurements and the prediction model, for a moving time horizon extending a fixed length into the past, as the following discrete time nonlinear MHE problem:

$$\min_{N, W} \sum_{k = -N_e}^{-1} \|W(k)\|_Q^2 + \sum_{k = -N_e}^{0} \|Y(k) - N(k)\|_R^2$$

subject to for $k = -N_e, \ldots, 0$:

$$Y(k) = \tilde{Y}(t - k)$$

for $k = -N_e, \ldots, -1$:

$$N(k + 1) = f_e(N(k), U(k), D(k), W(k)),$$

where $N_e$ is the estimation horizon, $k$ is the time step, $Q$ and $R$ are weighting matrices on demand and measurement noise, respectively, $\tilde{Y}(t)$ is the measurement taken at sampling instant $t$, $Y(k)$, $N(k)$, $U(k)$, $D(k)$, and $W(k)$ are the vectors containing all $Y_{IJ}(k)$, $N_{IJ}(k)$, $U_{IJ}(k)$, $D_{IJ}(k)$, and $W_{IJ}(k)$ terms, respectively, whereas $f_e$ is the time discretized version of equation (1).

**Model Predictive Control**

We formulate the problem of finding the $U_{IJ}$ values that minimize TTS as the following discrete time economic nonlinear MPC problem:

$$\min_{N, U} T \cdot \sum_{k = 0}^{N_p - 1} \|N(k)\|_1$$

subject to $N(0) = \tilde{N}(t)$

for $k = 0, \ldots, N_p - 1$:

$$N(k + 1) = f_p(N(k), U(k), D(k))$$

$$0 \leq N_{IJ}(k) \quad \forall I, J \in \mathcal{R}$$

$$N_I(k) \leq N_{jam}$$

$$U_{min} \leq U(k) \leq U_{min},$$

where $k$ and $T$ are the time step and sample time, respectively, $N_p$ is the prediction horizon, $t$ is the current sampling instant in time and $\tilde{N}(t)$ is the state estimate computed by the MHE at
that instant, $N(k)$, $U(k)$, and $D(k)$ are vectors containing all $N_{i,j}(k)$, $U_{i,j}(k)$, and $D_{i,j}(k)$ terms, respectively. $f_p$ is the time discretized version of equation (1) (with the assumption that all $W_{i,j}$ terms are 0, as these have zero mean but are unknown to the MPC), whereas $U_{\text{min}}$ and $U_{\text{min}}$ are the bounds on the perimeter control inputs.

**Integrated Moving Horizon Estimation and Model Predictive Control**

For the combined accumulation state estimation and perimeter control of large-scale urban networks, we propose a scheme integrating MHE and MPC (see figure 2). In this scheme, the MHE has access to information on noisy measurements $Y_{i,j}$ of accumulation states $N_{i,j}$, perimeter control inputs $U_{i,j}$, and average inflow demands $D_{i,j}$, for a fixed time horizon (i.e., $N_e$) into the past. Using these, at time $t$, the MHE computes the accumulation state estimate $\tilde{N}_{i,j}(t)$ by solving the problem (6). The state estimate is then used by the MPC, together with information on average inflow demand profiles $D_{i,j}$ for a fixed time horizon (i.e., $N_p$) into the future, to compute the perimeter control inputs $U_{i,j}(t)$ via solving the problem (7). The control inputs are applied to the urban network, completing the feedback loop.

**CONCLUSION**

In this paper we propose a combined estimation and control scheme based on real-time optimization, employing the MHE and MPC techniques, for the management of heterogeneously congested large-scale urban road networks. The full paper will include detailed simulation studies for evaluating the estimation and control performance of the proposed scheme for congested scenarios and different levels of uncertainty in the demands and measurements. Furthermore, the case studies will also include cases where the simulation model (i.e., the plant representing the reality) is a more detailed description of the traffic dynamics. To this end, a recently developed dynamical model (Lamotte and Geroliminis, 2017) will be considered, which is based on (in contrast to the MFDs network model with constant average trip lengths employed in this paper) trip length distributions. Evaluating the MHE-MPC scheme on more detailed traffic models is expected to shed more light on its value for practical applications.

**REFERENCES**


